



The Open
University

Science: a second level course

***Astronomy
and
Planetary Science***

Book 2

The Planets

Book chair: Dave Rothery

*Prepared for the Course Team by
Peter Francis, Charlie Harding, Elaine Moore, Colin Pillinger,
Dave Rothery, Ian Wright*

The Open University

S281 Course Team

Course Team Chair and General Editor Barrie Jones

Block 1 Chair Barrie Jones

Block 2 Chair Dave Rothery

Block 3 Chairs Barrie Jones and Bob Lambourne

Block 4 Chair Russell Stannard

Course Manager Cheryl Newport

Dave Adams *University of Leicester* (Author)

Jocelyn Bell Burnell (Author)

Cameron Balbirnie (BBC Producer)

Giles Clark (Publishing)

Alan Cooper (AV Production)

Sue Dobson (Graphic Artist)

Carol Forward (Course Secretary)

Peter Francis (Author)

John Greenwood (Library)

Charlie Harding (Author)

Karen Hill (Author)

Jonathan Hunt (Publishing)

Tony Jolly (BBC Series Producer)

Barrie Jones (Author)

Robert Lambourne (Author)

Jean McCloughry (Staff Tutor)

Elaine Moore (Author)

Lesley Passey (Designer)

Colin Pillinger (Author)

Ian Robson *University of Central Lancashire* (Author)

Dave Rothery (Author)

Dick Sharp (Editor)

Russell Stannard (Author)

Liz Swinbank (Consultant)

Margaret Swithenby (Editor)

Arnold Wolfendale *University of Durham* (Course Assessor)

Ian Wright (Author)

John Zarnecki *University of Kent* (Author)

Cover: Triton, the largest satellite of Neptune, as seen by the spaceprobe Voyager 2. Triton has a rugged icy surface, sculpted by icy volcanic processes, and a polar cap of frozen nitrogen. Image courtesy of the USGS Branch of Astrogeology, Flagstaff, Arizona.

The Open University, Walton Hall, Milton Keynes, MK7 6AA.

First published 1994

Copyright © 1994 The Open University.

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system or transmitted in any form or by any means, without written permission from the publisher or a licence from the Copyright Licensing Agency Limited. Details of such licences (for reprographic reproduction) may be obtained from the Copyright Licensing Agency Ltd of 90 Tottenham Court Road, London, W1P 9HE.

Edited, designed and typeset by The Open University.

Printed in Great Britain by Henry Ling Ltd, The Dorset Press, Dorchester, DT1 1HD.

ISBN 0 7492 51263

This text forms part of an Open University Second Level Course. If you would like a copy of *Studying with The Open University*, please write to the Central Enquiry Service, PO Box 200, The Open University, Walton Hall, Milton Keynes, MK7 6YZ. If you have not enrolled on the Course and would like to buy this or other Open University material, please write to Open University Educational Enterprises Ltd, 12 Coffieridge Close, Stony Stratford, Milton Keynes, MK11 1BY, United Kingdom.

1.1

The planets

Contents

Introduction and study guide for Block 2	4
Chapter 1 The Solar System	9
Chapter 2 Origin of the planets	26
Chapter 3 Layering of planets	47
Chapter 4 Cratering	68
Chapter 5 Volcanism	94
Chapter 6 Atmospheres of terrestrial planets	120
Chapter 7 The giant planets	168
Chapter 8 Meteorites and comets	204
ITQ answers and comments	240
SAQ answers and comments	249
Acknowledgements	259
Index	261

Introduction and study guide for Block 2

In Block 2, *The planets*, you have the opportunity to get to know your own backyard. The main aim is to examine the planets and other bodies that make up the Solar System and to understand some of the processes that occur upon and within them. Why does the Sun have planets? What are they made of? Why are some planets very different from others? Can we observe similar processes occurring on different planetary bodies? These are some of the questions you should be considering while you study this Block.

Chapter 1 is essentially a geography lesson, and introduces you to the size and layout of the Solar System. In it you will discover the different kinds of body that are present.

In Chapter 2 you will study the processes by which we think the Solar System, and in particular the planets and their satellites, came into being. Chapter 3 takes the story further by considering how a planet-sized ball of matter evolves once it has formed. This sets the scene for Chapters 4 and 5 where you will examine in detail just two of the processes that affect planetary bodies with hard surfaces; cratering in Chapter 4 and volcanism in Chapter 5.

Chapter 6 considers the formation, circulation and chemistry of planetary atmospheres, in particular those of the Earth, Venus and Mars. You will also learn how an atmosphere interacts with the surface material of a planet, and discover the significance of this for planetary processes and evolution.

In Chapter 7 you will study the four giant planets, and consider how their internal structure is liable to differ from that of the smaller bodies you examined in Chapter 3, and indeed from the stellar interiors that you considered in Block 1. You will also see how the driving forces for atmospheric circulation in these planets compare with those operating on the inner planets, and examine some of the exotic photochemical reactions thought to occur in the atmospheres of Jupiter, and Saturn's large satellite, Titan.

Finally, in Chapter 8, we return to the theme of the age, composition and origin of the Solar System, and look in some detail at what we can learn from the study of the primitive material preserved in meteorites and comets. By this stage you should be in a position to weigh the value of different types of information about Solar System material, planetary evolution and surface-atmosphere interactions, so we ask you to assess the evidence concerning the origin of an unusual meteorite, which may have come from Mars.

There are three TV programmes relevant to this Block. TV programme 3, *Cosmic recycling*, looks at how infrared telescopes help us to investigate star formation, and then examines some of the evidence that grains of interstellar origin have survived this process and been incorporated in meteorites that can be studied in the laboratory. TV programme 4, *Venus unveiled*, shows how close-up views from spacecraft have given us new insights into the processes that occur upon and within Venus. TV programme 5, *Design for an alien world*, assesses the current evidence as to the nature of Titan's atmosphere and shows how this is being used to design an instrument package intended to arrive on the surface of Titan in the year 2004.

There are also several video sequences associated with Block 2. Video sequence 5, *Tidal heating*, demonstrates the combination of orbital and rotational motion responsible for tidal heating in some of the satellites of the outer planets, and is also intended to give you a feel for how images such as many of those in *Images of the Cosmos* have been enhanced. Video sequence 6, *Cratering*, shows high-speed photography of experiments designed to produce craters by firing projectiles at a target surface, from which it can be inferred that the vast majority of craters we see on planetary surfaces are the result of impacts. Video sequence 7, *Atmospheric circulation*, shows time-lapse sequences from spaceprobes that reveal the contrasting patterns of atmospheric circulation on the giant planets,

and allows you to make some measurements for yourself. Video sequence 8, *Mass spectrometry of meteorites*, takes you on a brief conducted tour of the extremely sensitive, high-precision analytical facilities operated by the Planetary Science Unit within the Open University Department of Earth Sciences, which are used to analyse samples of meteorites and lunar rocks.

This Block is also supported by audiocassette. Band 2, *Planetary images*, takes you on a guided tour of Solar System bodies, as recorded on some of the pictures in *Images of the Cosmos*. In this sequence you are introduced to some of the fascinating processes that can be seen on the surfaces of planetary bodies, including some that we do not have space to discuss elsewhere.

It is important that you read the associated notes before you watch any of the TV programmes or video sequences, or listen to the audio band.

In your *Project file*, the project sheet entitled *The Moon* invites you to look at the surface of the nearest other planetary body to the Earth, the Moon, to see for yourself some of the evidence for cratering and volcanism discussed in Chapters 4 and 5. It would be best if you can do this before you reach Chapter 4, so if you have not yet begun this project, you should do so as soon as possible.

There are a vast number of bodies in the Solar System. You will be introduced to the main ones and how they fit into the scheme of things in Chapter 1. Some Solar System reference information is provided in Tables 1.1a–d. You are *not* expected to memorize any of this, but you may find it useful to refer to, especially in answering several of the ITQs and SAQs in this book, and probably also for TMAs and CMAs.

Much of our detailed knowledge of Solar System bodies has been obtained by spaceprobes that have flown past, gone into orbit round, or (rarely) landed on them. In this text we have chosen to concentrate on the basic data and the processes that these reveal, at the expense of discussing spaceprobes themselves. Some of the highlights of our exploration of the Solar System are listed in a table in *Images of the Cosmos*. This too is for reference only. You will learn more about some of these missions in TV programmes 4 and 5.

Table 1.1a Basic data on the planets (including the Moon). The mass and density of Mars have been left blank, because you will calculate these for yourself in ITQ 1.2.

	Mercury	Venus	Earth	Moon	Mars	Jupiter	Saturn	Uranus	Neptune	Pluto
Distance from Sun										
/ 10^6 km	57.9	108.2	149.6	(149.6)	227.9	778.3	1 427	2 870	4 497	5 900
/ AU ^a	0.38	0.72	1	(1)	1.52	5.2	9.5	19.1	30.0	39.4
Orbital period ^b	88.0	224.7	365.3	(27.3	687	11.86	29.46	84.01	164.79	247.69
days	days	days	days	days)	days	yr	yr	yr	yr	yr
Axial rotation ^b period / days	58.64	243.0	0.997	27.32	1.026	0.410	0.444	0.718	0.768	6.387
Axial inclination	0.0°	177.3°	23.4°	6.7°	2.2°	3.1°	26.7°	97.9°	29.6°	94°
Orbital inclination ^c	7.0°	3.4°	0.0°	5.2°	1.8°	1.3	2.5°	0.8°	1.8°	17.2°
Orbital eccentricity ^d	0.206	0.007	0.017	0.055	0.093	0.048	0.055	0.048	0.010	0.249
Equatorial radius / 10^3 km	2.439	6.052	6.378	1.738	3.394	71.4	60.0	25.6	24.3	1.18
Mass										
/ 10^{24} kg	0.330	4.87	5.98	0.0735		1900	569	86.8	102	0.014
/ M_{\oplus} ^e	0.0553	0.815	1.00	0.0123		317.8	95.2	14.5	17.2	0.002
Density / 10^3 kg m ⁻³	5.43	5.25	5.52	3.34		1.33	0.69	1.29	1.64	2.1
Surface gravity / m s ⁻²	3.7	8.9	9.8	1.6	3.7	24.8 ^f	10.6 ^f	8.9 ^f	11.6 ^f	0.4
Satellites	0	0	1	—	2	16	18	15	8	1
Rings	0	0	0	0	0	few	many	several	few	0?
Mean surface temperature / °C	170	460	15	1	-50	-143 ^f	-195 ^f	-201 ^f	-220 ^f	-205 to -165 ^g
Atmosphere (main components)	He Na O	CO ₂	N ₂ O ₂	Ne Ar H ₂ He	CO ₂	H ₂ ^f He CH ₄	H ₂ ^f He CH ₄	H ₂ ^f He	H ₂ ^f He	CH ₄ ?N ₂ ?CO
Mean surface pressure of atmosphere / bar ^h	10^{-15}	92	1.0	2×10^{-14}	0.006	n/a	n/a	n/a	n/a	variable
Mean surface density of atmosphere / kg m ⁻³	10^{-13}	67	1.225	—	1.8×10^{-2}	—	—	—	—	variable
Column mass of ⁱ atmosphere / kg m ⁻²	10^{-11}	1.1×10^6	1.0×10^4	10^{-11}	1.7×10^2	—	—	—	—	variable

^a AU stands for astronomical unit, and is defined as the mean distance of the Earth from the Sun.

^b These are sidereal periods. Days and years used in this table are Earth *solar* days and Earth years.

^c Orbital inclination is the inclination of the orbit to the ecliptic plane – the plane of the Earth's orbit about the Sun, used as a reference against which to measure the inclination of other orbits.

^d Eccentricity is a measure of the shape of an orbit. An orbit with an eccentricity of zero would be a perfect circle. The larger the value of eccentricity, the more strongly elliptical the shape.

^e The symbol M_{\oplus} stands for mass of the Earth, in the same way that M_{\odot} (which you met in Book 1) stands for mass of the Sun.

^f The four giant planets (Jupiter, Saturn, Uranus and Neptune) are fluid to considerable depth and have no definable surface; values of gravity, temperature and composition quoted here are for the atmospheric layer where the pressure is equal to the Earth's atmospheric pressure at sea-level.

^g The mean temperature of Pluto is poorly-known, but probably varies considerably as it moves around its eccentric orbit.

^h In this book we mostly use bar as our standard unit of pressure. 1 bar is close to the average atmospheric pressure on the Earth at sea-level, and is equivalent to 10^5 Pa.

ⁱ Column mass is the mass of atmosphere situated above each 1 m^2 area of the planet's surface.

Table 1.1b Planetary satellites

Planet	Satellite	Average distance from planet/ 10^3 km	Orbital period/ days	Equatorial radius/ km	Mass/ 10^{20} kg	Density/ 10^3 kg m $^{-3}$
Earth	Moon	384.5	27.32	1738	734.9	3.34
Mars	Phobos	9.4	0.319	14×10	1.3×10^{-4}	2.0
	Deimos	23.5	1.26	8×6	1.8×10^{-5}	1.7
Jupiter	Io	421.6	1.77	1815	894	3.57
	Europa	670.9	3.55	1569	480	2.97
	Ganymede	1070	7.16	2631	1482	1.94
	Callisto	1883	16.7	2400	1077	1.86
	12 others			< 135		
Saturn	Mimas	186	0.942	197	0.38	1.17
	Enceladus	238	1.37	251	0.8	1.24
	Tethys	295	1.89	524	7.6	1.26
	Dione	377	2.74	559	10.5	1.44
	Rhea	527	4.52	764	24.9	1.33
	Titan	1222	15.94	2575	1346	1.88
	Iapetus	3561	79.3	718	18.8	1.21
	11 others			< 175		
Uranus	Miranda	130	1.42	236	0.75	1.35
	Ariel	191	2.52	579	13.5	1.66
	Umbriel	266	4.15	586	12.7	1.51
	Titania	436	8.70	790	34.8	1.68
	Oberon	583	13.46	762	29.2	1.58
	10 others			< 85		
Neptune	Proteus	178	1.12	209	0.4	1.1
	Triton	345	5.88	1350	214	2.08
	Nereid	5510	365.2	170	0.27	1.27
	5 others			< 100		
Pluto	Charon	19.6	6.39	620	11	1.3

Table 1.1c The four largest asteroids

Asteroid	Average distance from Sun/ 10^6 km	Average distance from Sun / AU	Orbital period/yr	Orbital inclination	Orbital eccentricity	Equatorial radius / km	Mass/ 10^{20} kg
Ceres	413.9	2.77	4.61	10.6°	0.097	457	≈ 10
Pallas	414.5	2.77	4.61	34.8°	0.180	261	≈ 2
Vesta	353.4	2.36	3.63	7.2°	0.097	250	≈ 2
Hygiea	470.3	3.14	5.59	3.84°	0.136	215	≈ 1

Table 1.1d Selected comets

Comet	Perihelion ^a distance from Sun/ 10^6 km	Perihelion ^a distance from Sun / AU	Orbital period / yr	Orbital inclination	Orbital eccentricity
Bennett	80.5	0.538	$\gg 1000$	90.0°	0.996
Encke	51.0	0.341	3.31	11.9°	0.846
Giacobini-Zinner	154.7	1.034	6.61	31.8°	0.706
Halley	87.8	0.587	76.0	162.2°	0.967
Kohoutek	21.2	0.142	$\gg 1000$	14.3°	1.000^b
Swift-Tuttle	144.1	0.963	120	113.6°	0.960

^a The term perihelion means the closest point to the Sun reached during an orbit.

^b An eccentricity of 1.0 means that the orbit is a parabola (which is an infinitely long ellipse).

Chapter 1

The Solar System

Prepared for the Course Team by Peter Francis

Contents

1.1	Introduction	10
1.2	Layout of the Solar System	10
1.2.1	Kepler's Laws	11
1.2.2	Distances, sizes, masses and densities of Solar System bodies	13
1.3	Asteroids and comets: vermin of the sky	16
1.3.1	Asteroid types	18
1.3.2	Comets	19
1.4	Atmospheres	22
	Summary of Chapter 1 and SAQs	24
	Objectives for Chapter 1	24

1.1 Introduction

This chapter is a very small overview of something very large: the architecture of the Solar System. Astronomy is notorious for its statistics – the original ‘astronomical numbers’ with strings of zeros. In this chapter, we get to grips with some of these large numbers, and show that they are quite manageable when handled in a suitable context. Kepler’s laws of orbital motion are introduced, and you will use them yourself to work out the orbital periods of planetary bodies at different distances from the Sun. By considering their sizes, masses and densities, we show that there are several groups of planetary bodies in the Solar System, and that their properties and movements are intelligible in down-to-Earth ways, even if they are located on the far fringes of the Solar System. Of the smaller planetary bodies only a few possess an atmosphere. We show that even this attribute can be predicted from simple considerations of size, mass and temperature.

1.2 Layout of the Solar System

Living as we do near the start of the 21st century, we have less direct experience of the Universe than any previous generation. For millenia, the constellations visible in the night sky were familiar, and played a role in spiritual life. Naturally, the Sun and Moon dominated thinking about the Earth’s place in the cosmos, but five planets (Mercury, Venus, Mars, Jupiter and Saturn) were well known to the ancients because of their brightness and rapid motions relative to the ‘fixed’ stars. Apart from their astrological roles, the need for calendar systems led to sophisticated knowledge of planetary motions in several early cultures. Records of naked-eye observations of the motions of the planets against the starry background inspired the Polish astronomer Nicholas Copernicus (1473–1543) to realize in 1512 that the Earth revolves around the Sun, and the German astronomer Johannes Kepler (1571–1630) to deduce in 1595 that the planets move in elliptical, not circular, orbits. Kepler’s work led within 90 years to Sir Isaac Newton’s understanding of gravity and the laws of motion. Thus, simple visual observations of the motions of the planets underlie the whole of modern science. It was not until the Italian scientist Galileo Galilei (1564–1642) discovered four satellites of Jupiter in 1610 with a telescope that the existence of other bodies in the Solar System was known. Galileo’s observation that these satellites revolved around Jupiter also provided convincing proof that heavenly bodies can revolve around objects other than the Earth.

Since the invention of the telescope, the number of known bodies in the Solar System has increased enormously. Technological progress has led to occasional surges in that number. Construction of magnificent telescopes during the 19th and early 20th centuries led to the discovery of many planetary satellites and innumerable asteroids. In the two decades between 1970 and 1990, spacecraft explorations of the planets revealed yet more satellites and entire planetary ring systems. Where once only five planets were known beyond the Earth, it is impossible to state today how many objects there are, without reference to some minimum size. Every year, between 150 and 200 new asteroids are discovered, the smallest of them only about a hundred metres across. In the past, such minor objects were ignored by astronomers, who regarded them as nuisances, irrelevant to their investigations of the profound problems of space and time. Because of their relatively rapid motions, they left long, unsightly streaks across photographic plates of stellar subjects, earning for themselves the derisive term ‘vermin of the sky’. Today, however, study of minor bodies has blossomed, since in their diversity and antiquity, they provide rich sources of information on the early history of the Solar System.



Figure 1.1 (a) A diagrammatic summary of the Solar System, showing that most of the planets orbit the Sun close to the same plane as that of the Earth's orbit. Note that this is an oblique view – the orbits are actually nearly circular. The orbits of the five innermost planets are shown. (b) illustrates the relative sizes of the planets and the Sun.

In this section, we review briefly the organization of the Solar System, to provide a basis for later, more detailed sections that focus on individual processes and problems. Figure 1.1 illustrates graphically the sizes and positions of the major bodies, and Tables 1.1a–d provide statistics that may be useful at several points within the Block.

1.2.1 Kepler's laws

Kepler initially believed that the orbits of *all* celestial bodies had to be circular – anything else would represent a philosophical imperfection. In trying to reconcile actual observations of the movements of Mars with his notions of how he thought the Solar System should be constructed, Kepler realized that the orbit of Mars could *not* be circular – an early illustration of the 'scientific' method: using observational data to test an hypothesis. From this starting point, Kepler went on to formulate three laws of planetary motion, which remain fundamental to understanding the clockwork of the Solar System. They apply not only to the movements of the planets around the Sun, but also to the orbits of satellites – natural and artificial – around planets, and, in fact, to *all* orbital motions of a small body around a much larger one. Kepler was working at a time when the absolute distances of the planets from the Sun and from the Earth were unknown, so he worked entirely in relative terms, using the average distance of the Earth from the Sun as a reference. This later became known as the **astronomical unit (AU)**, and it is still a useful measure. **Kepler's first law states that:**

Planets move in elliptical orbits, with the Sun at one focus.

Recall from Book 1, Subsection 2.2.2, that the AU is now known to be approximately 150 million kilometres.

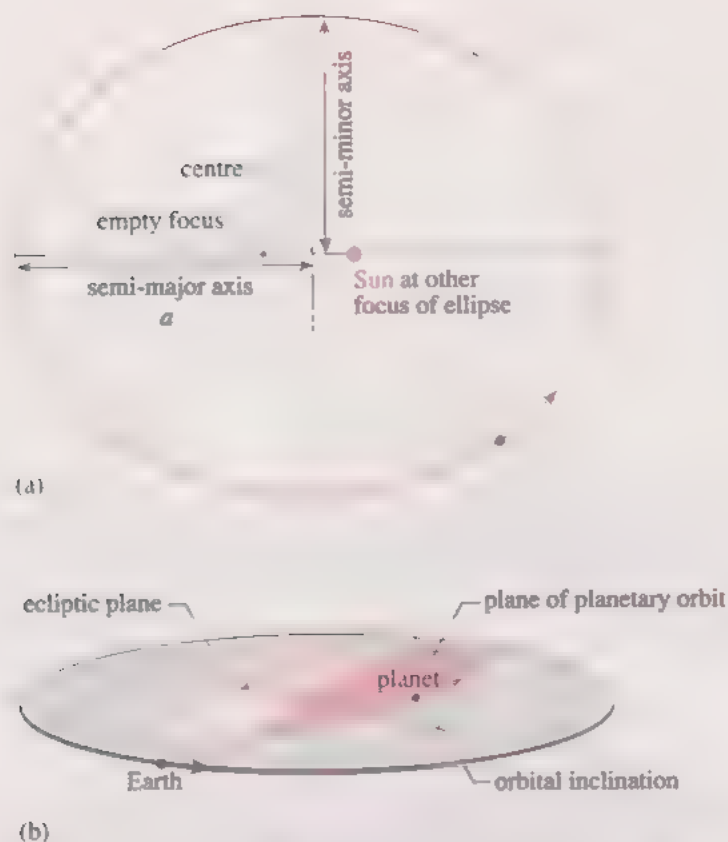


Figure 1.2 (a) An elliptical planetary orbit. Most of the planet's orbits are actually so nearly circular that it is difficult to show the ellipticity. Distances of planets from the Sun are given as the **semi-major axis** of the ellipse, which is the mean distance (the average distance). The Sun is not actually located precisely at the centre of the ellipse, but at a *focus*, whose position is described mathematically. In practice, the Sun is extremely close to the centre. The second focus is empty. The shape, or degree of flattening, of the ellipse, is expressed by the orbital **eccentricity**, which is the ratio of the distance from focus to centre, to the semi-major axis, a (cf. Figure 2.28 of Book 1). (b) Orbital inclination. The Earth's orbit around the Sun is shown as a plane, intersected by the plane of the orbit of a planet closer to the Sun. The orbital inclination of the planet is the angle between the two planes. For planets beyond the Earth, the plane of the Earth's orbit is imagined to be infinitely extended.

An elliptical orbit is shown in Figure 1.2a: you previously met elliptical orbits in Subsection 2.5.1 of Book 1. In fact, the orbits of the planets deviate only slightly from circular – Kepler was helped to his realization by the fact that Mars's orbit is the most elliptical of all the easily visible planets.

All the planets orbit the Sun in nearly the same plane, and move in the same direction. Two parameters are used to specify an elliptical orbit. The eccentricity (degree of flattening) is termed **orbital eccentricity**. The angle that the plane defined by this ellipse makes with the **ecliptic plane** (the plane of the Earth's orbit) is the **orbital inclination**, shown in Figure 1.2b. (The ecliptic plane is so called because eclipses only take place when the Earth, Moon and Sun are lined up in this plane.)

Which two planets in Table 1.1a are oddballs in terms of both orbital eccentricity and inclination?

■ The innermost planet (Mercury) and the outermost (Pluto) both have much higher eccentricities and inclinations than the others.

Pluto's orbit is so elliptical that it is sometimes nearer the Sun than Neptune – an inconvenience for compilers of general knowledge quizzes.

Kepler's second law is less obvious than his first. It states that:

The radius of a planet in its orbit sweeps out equal areas in equal times.

Figure 1.3 shows two such equal areas. Study this figure, and think about what the implications of the second law are for the speed of a planet along different parts of its elliptical orbit.

How does the speed of a planet between A and B compare with its speed between X and Y?

■ If the shaded areas in Figure 1.3 are swept out in the same time, the planet must move less quickly from A to B than from X to Y.

It follows that a planet does not move with uniform speed along an elliptical orbit. As it approaches the Sun, it speeds up; as it recedes, it slows. This means that calculations of orbital speeds for planets and satellites are complex for bodies in elliptical orbits. However, in practical terms these calculations are of relevance mainly to astrophysicists and scientists planning spacecraft trajectories. For our purposes, **Kepler's third law** suffices. It states that:

The square of a planet's period of revolution around the Sun, the **orbital period** (τ), is proportional to the cube of its average distance from the Sun (a).

The average distance is the same as the semi-major axis of its orbit. We can write this relationship as:

$$\tau^2 = k a^3 \quad (1.1)$$

or

$$k = \tau^2 / a^3 \quad (1.2)$$

For the Earth, τ is one year, and a is 1 AU. k , the constant of proportionality, is therefore $1 \text{ yr}^2 \text{ AU}^{-3}$. (The numerical value of k depends on the units used, such as days or seconds for time; metres or AU for distance.) This provides a simple method of finding the orbital period of any object in the Solar System, given its distance from the Sun, and vice versa.

11Q11 What would be the orbital period of a planet at a mean distance of 0.75 AU from the Sun?

1.2.2 Distances, sizes, masses and densities of Solar System bodies

Even through quite large telescopes distant planets appear merely as tiny discs, while their satellites remain faint points of light. What, therefore, can we possibly learn about the physical properties of such distant objects? How far away are they? How large or massive are they? What are they made of? These are elementary questions, but they can be quite difficult to answer.

In Kepler's day, distances of planetary objects could be established only in *relative* terms, using his third law (Equation 1.1). Although there were many earlier approximations, the value of the astronomical unit was not satisfactorily established until 1931. This measurement, and numerous earlier attempts, relied on the measurement of trigonometric parallax, or the apparent displacement of an object relative to its background when observed from different points, in the same way that the distances to stars are estimated (Book 1, Subsection 2.2.2). In 1930–31, a concerted effort was made by collaborating observatories all around the world to observe the parallax of the asteroid Eros against the starry background. Eros was chosen because its highly eccentric orbit brings it quite

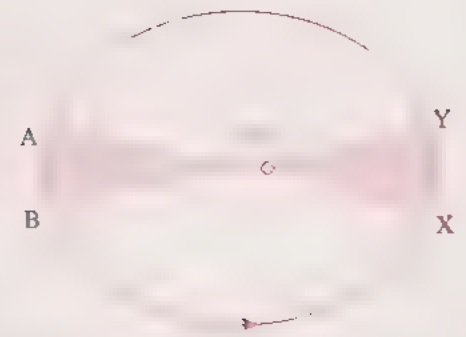


Figure 1.3 Kepler's second law states that the radius of a planet sweeps out equal areas in equal times. Two equal areas are shown shaded. The planet's speed along its orbit is not uniform, but varies with its distance from the Sun.

close to the Earth, when its parallax is easy to measure. Thirty telescopes at 24 observatories in 14 countries obtained 2 847 photographs, from which the absolute value of the semi-major axis of Eros's orbit could be derived. Thus using Kepler's third law, and the known orbital periods of Eros and the Earth, it was established that the mean Sun–Earth distance is 149.6 million kilometres.

Once the value of the astronomical unit had been established, it was easy to find the absolute distances of all the other objects in the Solar System using Kepler's third law. They are tabulated for the planets in Table 1.1a. *Do not just scan cursorily the numbers in this table.* Take a piece of string, and knot it at distances scaled to represent the distance of each planet, say 10 cm per 10^8 km. This will take 5–10 minutes, but it is the only way to gain an intuitive appreciation of how much further from the Sun the outer planets are, compared to those in the inner Solar System.

If you need 5.9 m of string to reach Pluto, how much would you need to reach the nearest star, 4.2 light years or 4×10^{13} km away?

■ About 40 km.

If a planetary body is big enough or close enough to be resolved into a disc with a telescope, and its distance is known, then estimation of its radius is straightforward. As you saw in Book 1, Subsection 2.5.1, if bodies of mass M and m are in orbit around each other, then by measuring the orbital period and semi-major axis we can get $(M + m)$. For a satellite in orbit around a planet, the mass M of the planet is so much greater than the mass m of the satellite that, to a good approximation, $(M + m)$ is equal to M , so we get the mass of the planet. It is much more difficult to determine the masses of satellites and asteroids, which do not have objects conveniently orbiting them. In recent years, the masses of many small planetary satellites (Table 1.1b) have been estimated by measuring the accelerations experienced by spacecraft whose trajectories carried them close by. These determinations provide a useful way of comparing the properties of the various objects in the Solar System.

HO 1.2 To demonstrate how straightforward it is to derive the mass of a planet from observations of its satellites, we now ask you to calculate the mass of Mars, deliberately left blank in Table 1.1a. Mars has two tiny satellites. One of these, Deimos, orbits Mars at a distance of 23 460 km with a period of 1.263 days (1 day = 86 400 s). You will need to use Equation 2.15 from Book 1:

$$M + m = \frac{4\pi^2 a^3}{G\tau^2}$$

remembering to work in SI units ($G = 6.67 \times 10^{-11} \text{ N m}^2 \text{ kg}^{-2}$). Because Deimos is so tiny (only about 14 km across), its mass is negligible. If this calculation goes well, go on to find Mars's density, given that its radius can be measured telescopically as 3 400 km. (Volume of a sphere = $\frac{4}{3}\pi R^3$.)

Perhaps you can now see where Kepler's third law comes from: in Equation 2.15 of Book 1, put M equal to the solar mass, and remember that planetary masses are small compared to the Sun's mass M_\odot . We get

$$\tau^2 = \left(\frac{4\pi^2}{GM_\odot} \right) a^3$$

So k in Equation 1.1 is equal to

$$\frac{4\pi^2}{GM_\odot}$$

Tables 1.1a, b and c list the masses, sizes and densities of the planets, their major satellites and some asteroids. It would be useful to explore the relationships between the sizes of bodies and their densities by displaying these data graphically.

Why would it be difficult to plot their radii on a graph in simple linear units?

■ The range in radius is so enormous (from about 7 km to more than 70 000 km), that the data would not all usefully fit on an ordinary linear graph – the smaller-sized objects would all be crammed up at one end.

The problem is overcome by using a graph with one 'ordinary' axis, for densities, and one logarithmic axis (see *Preparatory science*, Mathematical appendix) for radii. Such graphs are often termed 'log-linear'. On a logarithmic graph, each major interval is a power of ten – 10, 100, 1 000 and so on. You have seen in Book 1 that such graphs are a convenient way of plotting data that range across many powers of ten.

Use the graph in Figure 1.4 to plot a selection of planetary and satellite radii against the densities of the objects, writing the name of each object (abbreviated if necessary) alongside its position on the graph, to help you identify it.



Figure 1.4 Log-linear graph for plotting planetary radius and density.

11013 Do the objects cluster into any distinct groups? If so, what groups can you see?

Make sure that your plot matches that in Fig. 1.11 before reading on, and that you agree with the following summary.

The term **terrestrial planet** is defined differently according to context. The strictest definition includes only the four innermost planets (Mercury to Mars), but when used in a compositional sense it can include the Moon, Io and even Europa.

Four planets form an obvious family, characterized by huge dimensions (radii greater than 10 000 km) and low densities. These are the **giant planets**, dominated by Jupiter, whose mass is more than that of all the other planets put together. They must be composed mostly of low-density materials, which the cosmochemical arguments outlined in Chapters 2 and 7 dictate must be **dominantly hydrogen and helium**.

A second group straggles off towards the right, high-density side of the graph. It includes rather dense, but fairly large bodies, with radii of more than 1 000 km, notably the so called **terrestrial planets**, Mercury, Venus, Earth and Mars, and three dense satellites – our own Moon, plus Jupiter's satellites Io and Europa. Their densities suggest that these bodies are dominantly rocky, like the Earth.

The minerals that make up rock (about which you will learn more in Box 2.1) have densities between 2.5 and $3.5 \times 10^3 \text{ kg m}^{-3}$. Is it likely that the terrestrial planets could consist *exclusively* of rocky material?

■ No, they must contain a denser component.

This denser component is thought to be mostly metallic iron, but it may in some cases be iron sulphide.

The two tiny satellites of Mars – Phobos and Deimos – plot right at the bottom of the graph. They are probably 'captured' asteroids. Unfortunately, the sizes, masses and densities of most asteroids are too poorly known to plot them usefully on Figure 1.4. If more asteroid data were available, they would probably plot on a field extending up from Phobos and Deimos towards the terrestrial planets.

Clustered towards the left side of the graph are a large number of satellites, which form a third indistinct group, with radii of a few hundred to a few thousand kilometres, and densities between 1 and $2 \times 10^3 \text{ kg m}^{-3}$. These consist mostly of ice, with varying proportions of rocky materials. Low-density satellites such as Mimas must consist almost entirely of ice, whereas denser examples such as Ganymede and Callisto must contain substantial proportions of rock. There is no clear division between these icy objects and the terrestrial planets, and Europa represents a transitional case, having an icy layer a few tens of kilometres thick overlying its rocky interior.

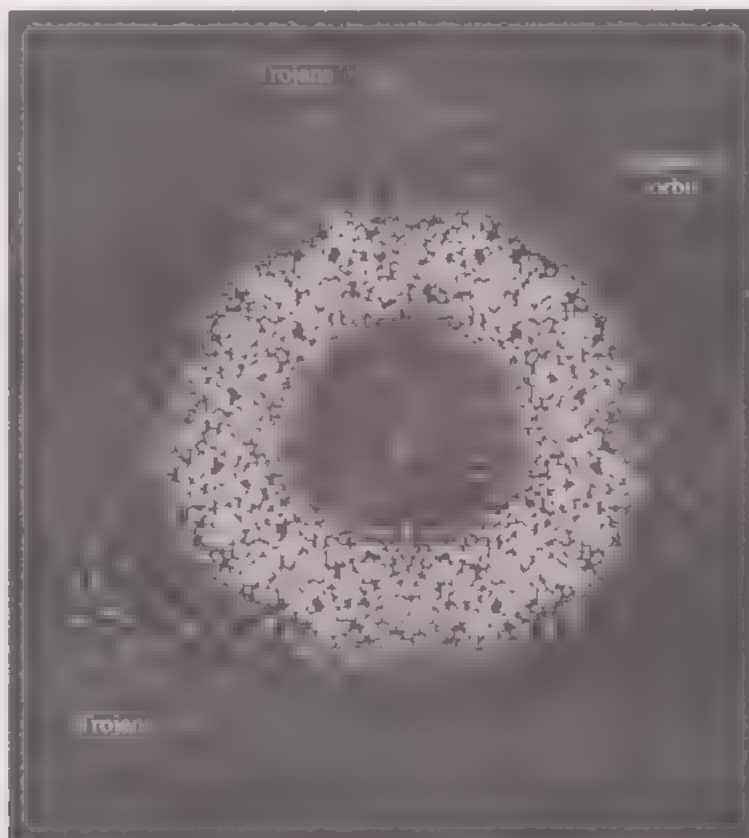
11Q 1.4 What does Pluto's position on the graph tell you about its nature? How does it compare with its neighbour, Neptune?

1.3 Asteroids and comets: vermin of the sky

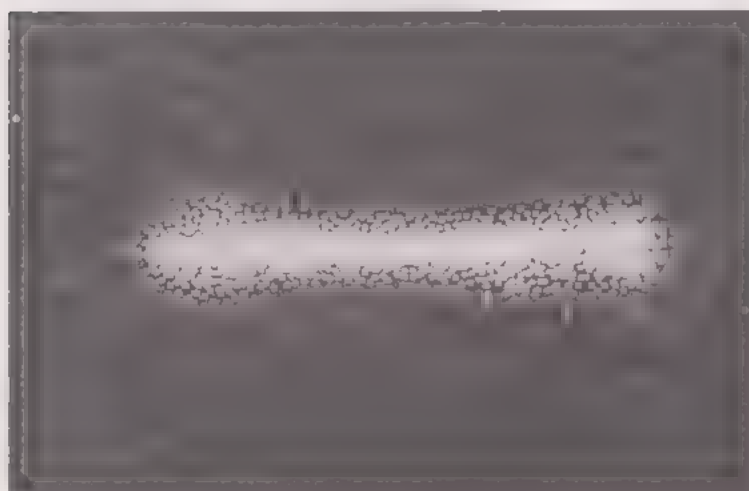
Asteroids are by far the most abundant named objects in the Solar System. Nearly five thousand are known, most occupying the 'asteroid belt' at 2.2–3.2 AU from the Sun, between the orbits of Mars and Jupiter (Figure 1.5a and b). The largest is Ceres (pronounced 'series'), with a radius of 457 km. Most are much smaller – only a few kilometres in diameter – and innumerable still smaller anonymous boulders probably orbit within the belt. Gaspra, visited by the Galileo spacecraft en route to Jupiter, is probably typical (Plate 2.43). 19th and 20th century astronomers concluded that the teeming asteroids were fragments of a single planet which had disintegrated catastrophically, perhaps due to massive gravitational perturbations caused by Jupiter. This imaginative view is no longer tenable, since although very numerous, the cumulative *mass* of all the asteroids is less than one-thousandth that of the Earth. Most asteroids are now thought to represent fragments of *many* small planetary bodies, each with its own unique characteristics, but many were never parts of larger objects. A few asteroids with **unusual orbits** may perhaps be remnants of comets.

Jupiter has, none the less, exerted a profound influence over the asteroid belt. When the numbers of asteroids at a given distance from the Sun are plotted against distance, a striking pattern emerges, as shown in Figure 1.6: there are spaces, or gaps, within the asteroid belt that contain few asteroids. These gaps, known as **Kirkwood gaps**, after their discoverer, are not random, but occur where there is **orbital resonance** between the orbital period of Jupiter and the period that an asteroid would have at that distance from the Sun, as given by Kepler's third law. For example, there is an important gap near 2.5 AU. At this distance, an asteroid would orbit the Sun exactly three times while Jupiter orbited once. This is a 3 : 1 resonance.

Similar gaps are seen in planetary ring systems. The origin of these elegant relationships involves subtle gravitational interactions between more and less massive objects. If a small object has an orbit which is resonant with a larger one,



(a)



(b)

Figure 1.5 (a) Nearly 5 000 individual asteroids are plotted here in their positions on 1 January 1995. Their distribution shows that the asteroid 'belt' is actually a diffuse cloud. Icarus and Apollo (not labelled) have Earth-crossing orbits. Note the locations of the two groups of Trojans, occupying what are called the Lagrangian points of Jupiter's orbit, 60° ahead of and 60° behind the planet. There may be several hundred Trojans at each Lagrangian point. (b) A cross-section through (a) (to the same scale) showing the thickness of the diffuse cloud of asteroids forming the asteroid belt. Each individual asteroid moves in an orbit inclined to the ecliptic plane, so that sometimes it is above it, and sometimes below.

there will be times when it is being accelerated forward in its orbit by the gravitational pull of the larger body, and other times when it is being decelerated. The cumulative effect of these forces is to distort the orbit of the smaller object, until it no longer has a resonant period, and its former orbit remains unoccupied. In the case of Saturn's beautiful ring system (Plate 2.51), the most prominent gap (known as Cassini's division) is defined by a resonance with the satellite Mimas: a particle orbiting on the inner edge of the division would have a period half that of Mimas.

11Q 1.5 Jupiter's period is 11.86 years. How far from the Sun, in AU, would you expect to find a gap in the asteroid belt corresponding to a $5:2$ resonance?

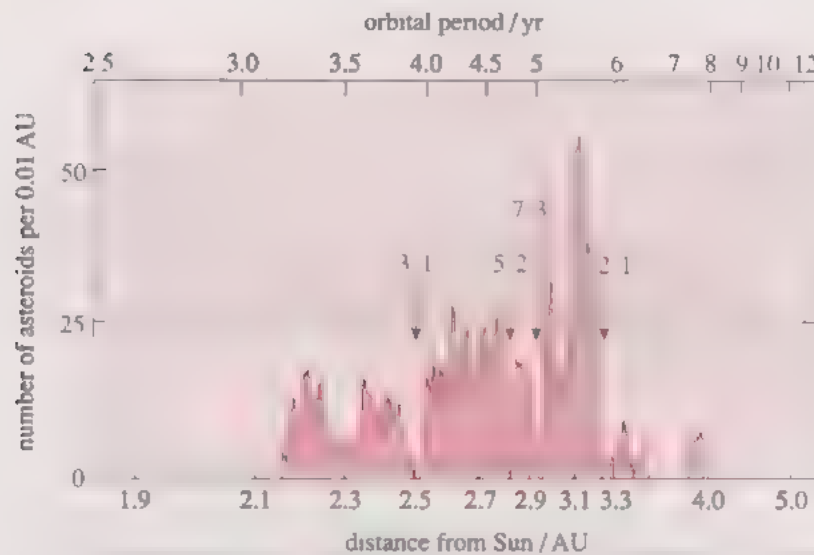


Figure 1.6 Variations in abundance of asteroids within the asteroid belt, showing the Kirkwood gaps at distances at which asteroids would be in orbital resonance with Jupiter.

These are known as Earth-crossing asteroids.

In scientific usage, 1 billion = 10^9 .

Some asteroids have such eccentric orbits that they cross the Earth's orbit, and thus potentially could collide with it. A collision which took place 65 million years ago may have caused the extinction of the dinosaurs, and there have been many lesser ones since. Only about 40 000 years ago, Meteor Crater in Arizona (USA) was formed by the impact of an object only 40–50 m in diameter.

Today, well over 4 billion years after the origin of the Solar System, there are still numerous asteroids which could collide with the Earth. The lifetimes of these asteroids must be short relative to the age of the Solar System, because they are rapidly removed by collisions. What does this imply?

■ The supply of Earth-crossing asteroids must somehow be replenished.

Current thinking suggests that fresh asteroids are ejected from the asteroid belt into Earth-crossing orbits from time to time, perhaps because of inter-asteroid collisions or orbital perturbations caused by Jupiter.

1.3.1 Asteroid types

In recent years, much effort has been devoted to mapping the distributions of asteroids of different types within the asteroid belt by measuring their brightness at different wavelengths – a measure of their 'colour'. C-type asteroids are rather dark, for example, whereas S-type asteroids are much brighter, and reddish. When the positions of asteroids of different classes are plotted, a pattern of sorts

emerges: S-type asteroids are most abundant in the inner part of the belt; C-types dominate further out (Figure 1.7).

Comparing the colour types of asteroids with those of meteorites has been fruitful, because it is difficult to determine the masses, densities and thus compositions of asteroids directly. **Meteorites** (discussed extensively in Chapter 8) are fragments of solid material that reach the Earth's surface from space; they are not to be confused with **meteors**, which are tiny dust-like grains burning up in the Earth's atmosphere to form 'shooting stars'. S-type asteroids have been found to correspond closely with meteorites consisting of mixtures of metallic iron and rocky material. C-type asteroids correspond with a group of carbon-rich rocky meteorites. For now, we can conclude that asteroids are small, dense rocky bodies, broadly similar in composition to the terrestrial planets. Studies of meteorites have had profound implications for theories on the origins of Solar System, discussed further in Sections 2.2–2.3. Perhaps their most important single result has been to provide the age of the Solar System, about 4.6 billion years (4.6×10^9 years).

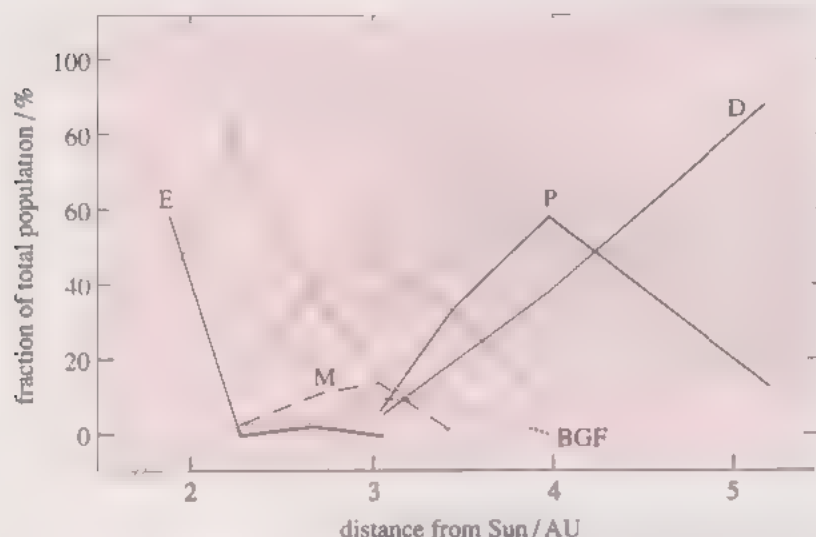


Figure 1.7 Distributions of some of the major types of asteroid within the asteroid belt. (Several different colour types shown. You will not be expected to recall all these.)

1.3.2 Comets

When beggars die there are no comets seen:

the heavens themselves blaze forth the death of princes.

This famous quotation from Shakespeare's *Julius Caesar* reminds us that **comets** have traditionally been thought of as portents of great events – a great comet was seen prior to Julius Caesar's assassination in 44 BC. The best-known comet of all is comet Halley, named after the noted British scientist Edmund Halley (1656–1742), a friend of Sir Isaac Newton. He showed that comets orbit the Sun as members of the Solar System, and are not stray visitors from elsewhere. This fact had not been appreciated previously, because of the long period between successive reappearances of most individual comets. Halley made observations of a spectacular comet seen in 1682, realized that it had previously been seen in 1531 and 1607; and, with help from Newton, worked out its orbit and correctly predicted its reappearance in 1758, 16 years after his death.

Figure 1.8 illustrates the difficulty Halley encountered in calculating the orbit of a comet: near the Sun, the shape of the orbit is close to that of a parabola, a mathematical figure that closely resembles an ellipse near its focus, but which, being infinitely extended, has no second focus. Comet Halley belongs to a group

which are called **short-period comets**; their orbits are elliptical, and they return relatively frequently. **Long-period comets**, by contrast, move on extended orbits which are so nearly parabolic near the Sun that even today it is difficult to determine their orbits precisely. Many are thought to have periods of thousands or even tens of thousands of years, their orbits carrying them to the most distant reaches of the Solar System. Comets Bennett and Kohoutek (pronounced 'ko-hoo-tek') in Table 1.1d are examples of this class. Unlike the planets, comets show a wide range in the *inclinations* of their orbits, so that many of them swoop through the ecliptic plane at large angles as they approach the Sun. Comet Halley's eagerly awaited reappearance in 1986 was disappointing, partly because its nearest approach to the Sun took place when the Earth was badly placed, on the other side of the Sun, and partly because light pollution now makes it difficult to see any but the brightest celestial objects. Fortunately, however, the Giotto spacecraft made a successful fly-by, providing the first-ever close-up images of any comet (Plate 2.45).

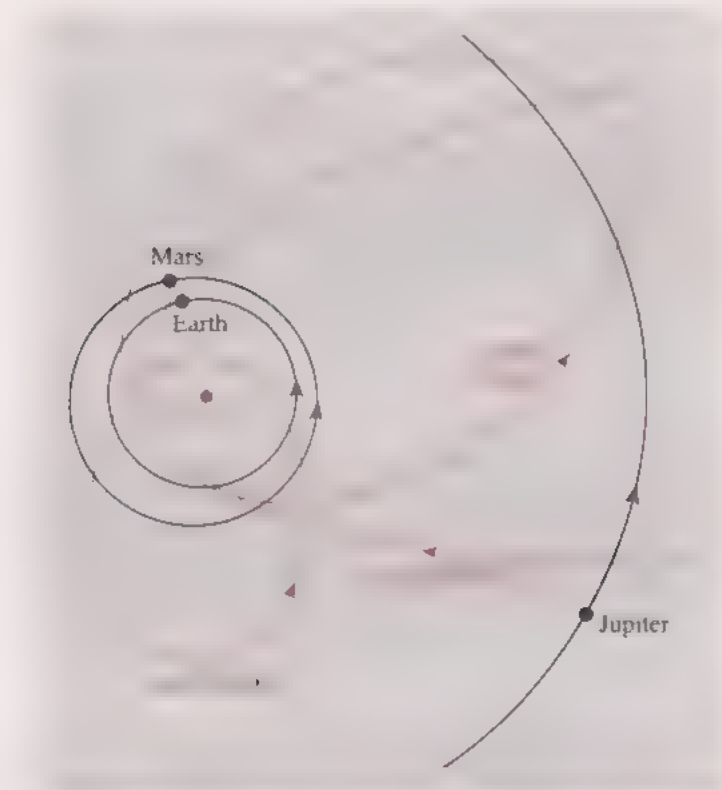


Figure 1.8 Orbits of some short-period comets. At its furthest from the Sun, the orbit of comet Halley carries it out beyond Neptune. These orbits are not in the ecliptic plane.

If a comet had a truly parabolic orbit, what would this imply about its origin?

■ Because a parabola is an open figure, such a comet could only be a visitor from deep space, beyond the Solar System. (Actually, most comets come from the Oort cloud; see later.)

Since the work of the American astronomer Fred L. Whipple (1906–) in 1950, comets have been thought of as 'dirty snowballs' – lumps of ice mixed with silicate particles. Far from the Sun, a comet is an inert lump. As it approaches the Sun, solar heating evaporates some ice containing fine dust particles, so a comet loses material in the formation of an expanding gaseous envelope and a long tail. A typical comet thus consists of three components: the key component is the compact **nucleus**, only a few kilometres across; surrounding the nucleus is a **coma**, an ephemeral atmosphere of tenuous gas and dust particles, from which

extends the long **tail** directed away from the Sun. Sunlight reflected from dust particles in the coma and tail makes them visible, while fluorescence of ionized gas also adds a glow. During the 1910 appearance of comet Halley, its magnificent tail streamed over no less than 40° of the night sky (Figure 1.9).

Comet tails extend over such large parts of the inner Solar System that planetary orbits inevitably intersect them. When the Earth passes through a comet tail, minute particles burn up in the atmosphere, but few meteorites which reach the ground seem to be derived from comets. Much more rarely, the nucleus of a comet hits the Earth. A minor event of this sort may have taken place at Tunguska in Siberia in 1908, when a violent explosion flattened more than a million square kilometres of forest.

Spacecraft encounters with comet Halley in 1986 supported the 'dirty snowball' hypothesis. They showed that the nucleus was actually a peanut-shaped object, about $16\text{ km} \times 8\text{ km}$ in size. Estimates of its density were surprisingly low, about $0.25 \times 10^3\text{ kg m}^{-3}$. Such a low density is consistent with an icy composition; the ice vaporizing near the Sun to provide the gaseous

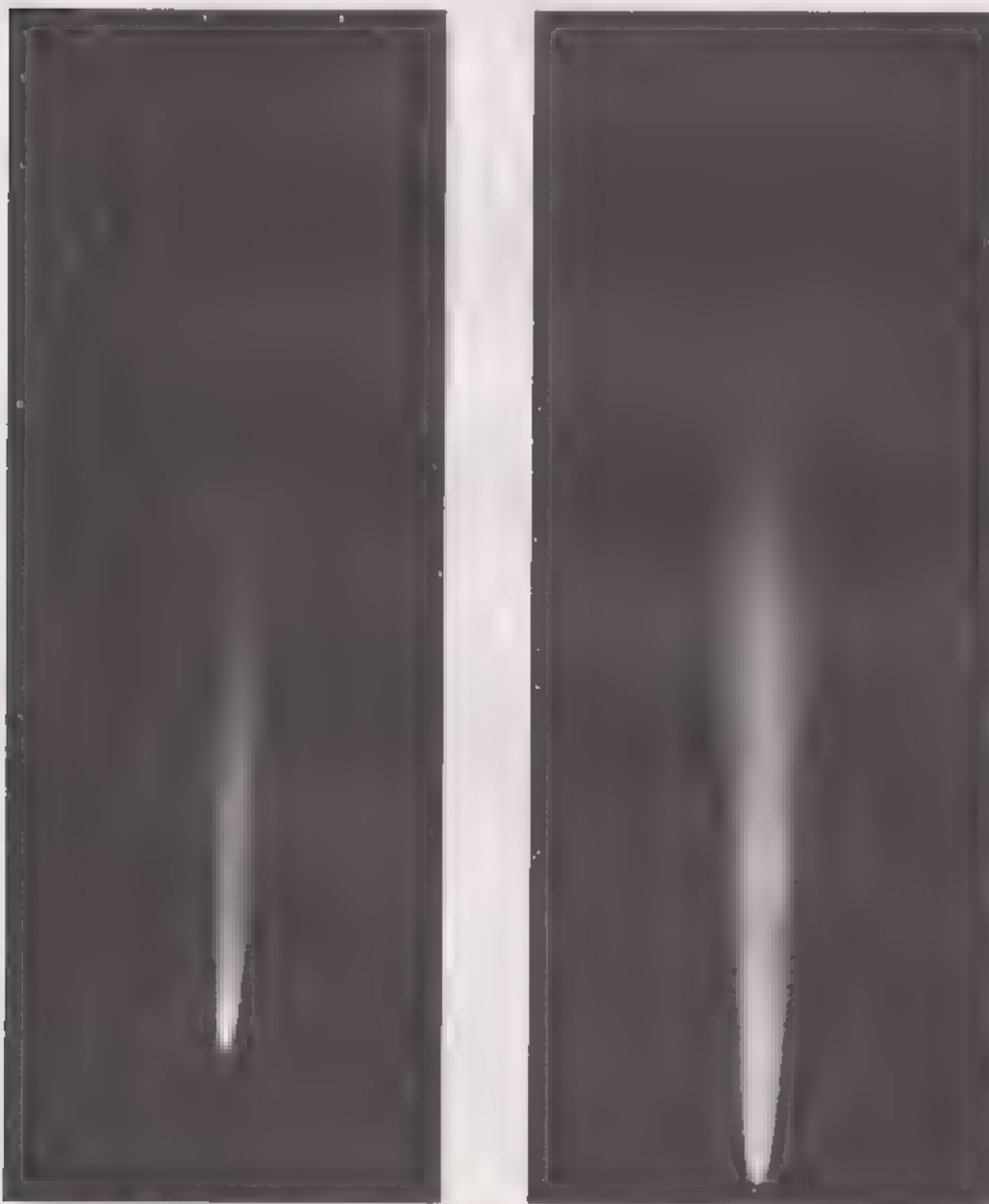


Figure 1.9 Two views of comet Halley during its spectacular apparition of 1910, when its tail extended up to 40° across the night sky. These photographs were taken in Honolulu, Hawaii

envelope of the coma and leaving cavities in the interior. On each of its passages past the Sun, a comet loses some fraction of its volatile icy material. Ultimately, it will all be lost, leaving only a rocky and carbonaceous residue. Some Earth-crossing asteroids have highly elliptical orbits similar to comets. It is possible that they are the burned-out husks of old comets. It is tempting to argue in such cases that all that is left of the original ‘dirty snowballs’ is the dirt.

A single comet, however inspiring to behold, is an insignificant object in terms of mass. What comets may lack individually, however, they make up for in numbers. Jan Oort (1900–1992), a Dutch astronomer, suggested in 1950 that most comets reside in an essentially spherical cloud surrounding the Sun at a distance of 20 000–100 000 AU. (The nearest stars, in the Alpha Centauri group, are about 270 000 AU, or 4 light years distant.) This **Oort cloud** may contain about 10^{12} comets, with a cumulative mass some 25 times that of the Earth. There is no direct observational evidence for the existence of so many comets, but most astrophysicists accept the merit of Oort’s statistical arguments.

1.4 Atmospheres

Most objects in the Solar System are small, airless lumps. The four giant planets have extensive gaseous envelopes of hydrogen and helium, overlying fluid interiors of similar composition. The larger solid objects, of which the Earth is one, have envelopes of gas which form atmospheres. Venus, Earth and Mars all have atmospheres dense enough to play significant roles in their surface processes. Venus’s atmospheric pressure is 90 times the Earth’s, while Mars’s is less than 1% that of the Earth. Of the planetary satellites, only Saturn’s satellite Titan (which is larger than the planet Mercury) has a dense atmosphere. Chapter 6 deals with planetary atmospheres at length, but it is appropriate to enquire here why some planetary bodies have atmospheres, and others don’t.

□ What, if anything, do Venus, Earth, Mars and Titan have in common?

■ Not much, except that they all fall within roughly the same size range (Figure 1.11). Titan is one of the largest satellites in the Solar System.

Ultimately, it is the strength of the gravitational field at its surface that dictates whether or not an object in the Solar System can retain an atmosphere: the stronger the field is, the stronger the gravitational forces acting on the molecules in the atmosphere. This leads to the notion of **escape velocity** (*Preparatory science*, Section 1.5), the smallest upward speed that any object (spacecraft or molecule) must have to escape from a body. In *Preparatory science* it was established that the escape speed for a body of mass M and radius R is given by,

$$v_{\text{esc}} = \sqrt{2GM/R} \quad (1.3)$$

Whether atmospheric molecules have sufficient speed depends on the temperature. As the temperature of a gas increases, its molecules move around more quickly, and the average speed of its molecules increases (*Preparatory science*, Subsection 2.4.1). Some fraction of the molecules will always be travelling fast enough to overcome gravitational forces, allowing them to escape to space. At low temperatures, this proportion is negligible, but at higher temperatures it becomes progressively more significant, until most molecules exceed the escape speed for the planetary body. Note that the relevant temperature is that at a level in the upper atmosphere above which the atmosphere is so thin that a molecule moving outwards has little chance of colliding with another, and so *will* escape if it has sufficient speed.

Different gases have different molecular masses, so their average speeds are different at a given temperature. In order for a **planetary body** to retain a particular gas in its atmosphere for a period of time of the same order as the age of the Solar System, the average speed of the molecules in the gas should be less than about one-sixth of the escape speed. (If the average speed exceeds one-sixth of the escape speed, a significant proportion of molecules will be moving faster, and will be lost.) This condition is achieved on only a few planets and satellites. Mercury is so close to the Sun and so hot that average molecular speeds for all common gases are too great. Titan, which is a similar size to Mercury, is much less dense ($\approx 1.9 \times 10^3 \text{ kg m}^{-3}$). This in turn means that its mass is less than half of Mercury's, and therefore its surface gravity and escape speed are lower. Titan, however, is also so far from the Sun that its temperature is only about 100 K. Thus, Titan can retain a dense atmosphere.

'Planetary body' is a handy term that can be used to encompass planets, satellites and asteroids

Figure 1.10 (a log-log graph) explores these relationships further. For each of the planets (including the Moon), the temperature is plotted along the horizontal axis and one-sixth of the corresponding escape speed on the vertical axis. For the named gases, the sloping lines show at each temperature the average molecular speeds of each gas (strictly the root mean square speeds – see *Preparatory science*, Subsection 2.4.1). Figure 1.10 thus defines the conditions under which a planet would lose or retain that gas over geological time-scales; i.e. thousands of millions of years. The giant planets plot well above all the lines: they can therefore retain any gas. The Moon plots below all the lines – it can retain no gases.

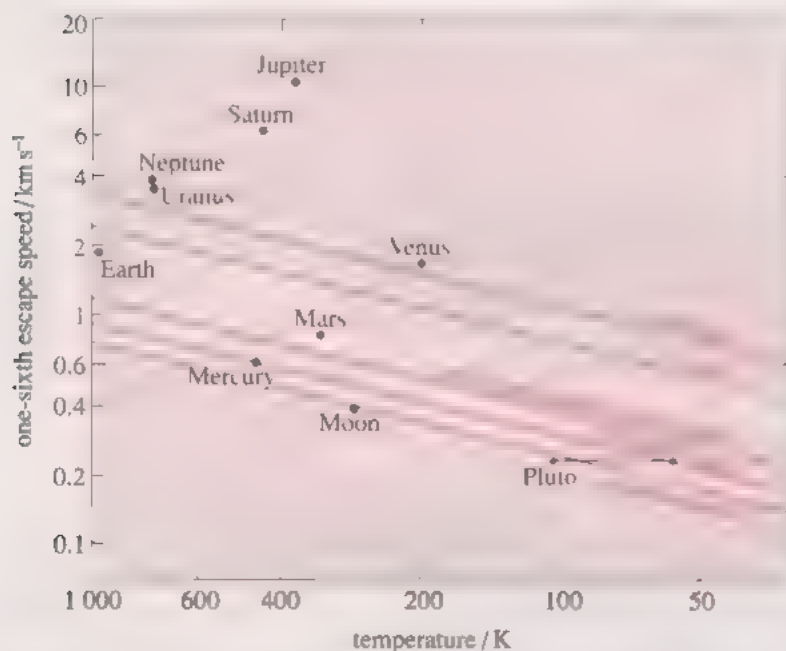


Figure 1.10 Graph summarizing conditions of temperature and escape speed for which planetary bodies can retain common gases in their atmospheres for long periods: hydrogen (H_2), helium (He), water vapour (H_2O), ammonia (NH_3), methane (CH_4), nitrogen (N_2), oxygen (O_2) and carbon dioxide (CO_2). For bodies with substantial atmosphere, the temperature is that at the top of the atmosphere. For other bodies it is the mean surface temperature. Pluto's temperature varies greatly according to its position around its orbit.

□ Which gases should Mars be able to retain?

■ Mars plots below the hydrogen and helium lines, so cannot retain these gases; but it plots above those for water vapour, nitrogen, oxygen, and carbon dioxide, so these gases could be retained. (These gases have actually been found on Mars.)

We recommend that you listen to audio band 2 before you begin the next chapter. This band introduces you to the selection of planetary images in Images of the Cosmos.

Summary of Chapter 1 and SAQs

In this chapter, we have briefly reviewed the architecture of the Solar System. We have seen that Kepler's laws of orbital motion can be used to predict the motions of *any* planetary body. The sizes, masses and densities of the main bodies in the Solar System fall into three indistinct groups: the giant planets, rocky bodies (including the terrestrial planets and ranging down in size through the asteroids), and the icy satellites. Most of the small bodies lack an atmosphere, but we have seen that the factors which dictate whether or not a planetary body retains an atmosphere are its surface gravity (which is in turn controlled by size and mass) and its surface (or upper atmospheric) temperature.

SAQ 1.1 (Objectives 1.2 and 1.3) In 1977 a small object, about 90 km in diameter, was found orbiting between Saturn and Uranus with a strongly elliptical orbit that brings it closest to the Sun at 8.5 AU (perihelion) and furthest at a distance of 18.9 AU (aphelion). It was named Chiron. Its semi-major axis is thus $((18.9 + 8.5)/2)$ AU = 13.7 AU. What is Chiron's orbital period? What class of objects is it most likely to resemble?

SAQ 1.2 (Objectives 1.3 and 1.4) Using the data given in Table 1.1c and the text, to which of the groups of planetary bodies that you located on Figure 1.4 do you think that (a) comet Halley and (b) the asteroid Vesta belong? Why would you expect Vesta to possess virtually no atmosphere?

SAQ 1.3 (Objective 1.2) The Earth orbits the Sun at a distance of 150 million km in 365 days. If it were transported into a 150 million km radius orbit about another star, would it necessarily orbit that alien sun in the same period?

Objectives for Chapter 1

After studying Chapter 1 (and any associated audio, video or TV material), you should be able to:

- 1.1 Give brief definitions of the terms, concepts and principles listed below.
- 1.2 Given the orbital period of a planetary body, calculate its mean distance from the Sun and vice versa.
- 1.3 Given data on the sizes, masses and densities of planetary bodies, predict their likely compositions and affinities within the Solar System.
- 1.4 Given appropriate size, mass and temperature data, predict the range of gases that a planetary body may be able to retain in its atmosphere.

List of scientific terms, concepts and principles used in Chapter 1

Term	Page	Term	Page	Term	Page
asteroid	16	Kirkwood gaps	17	orbital period	13
astronomical unit (AU)	11	long-period comets	20	orbital resonance	17
coma	20	meteorites	19	planetary body	23
comet	19	meteors	19	semi-major axis	12
ecliptic plane	12	nucleus (of a comet)	20	short-period comets	20
escape velocity	22	Oort cloud	22	tail (of a comet)	21
giant planet	16	orbital eccentricity	12	terrestrial planet	16
Kepler's first, second and third laws	11,13	orbital inclination	12		

Chapter 2

Origin of the planets

Prepared for the Course Team by Dave Rothery

Contents

2.1	Introduction	27
2.2	The solar nebula	27
2.2.1	Nebular theories	27
2.2.2	Physical processes in a collapsing cloud	28
2.2.3	The composition of the solar nebula	31
	Summary of Section 2.2 and SAQs	33
2.3	Condensation and coagulation	33
2.3.1	Grains in the solar nebula	34
2.3.2	Coagulation of grains	35
	Summary of Section 2.3 and SAQs	37
2.4	Planetesimal formation	38
2.5	Planetary embryos	38
2.5.1	Runaway growth	38
2.5.2	Planetary growth in the inner Solar System	39
2.5.3	Formation of the giant planets	40
2.5.4	Origin of the asteroids and comets	42
	Summary of Sections 2.4 and 2.5 and SAQs	43
2.6	Satellite systems	43
	Summary of Section 2.6 and SAQs	46
	Objectives for Chapter 2	46

2.1 Introduction

In Book 1, Subsection 3.2.3, you were introduced to some of the processes occurring when interstellar matter contracts and initiates star birth. In particular you saw how young stars, particularly pre-main sequence T Tauri stars, lie within clouds of molecular gas and dust representing remnant pre-stellar material, and it was suggested that planets could be forming within circumstellar discs around young stars.

The dust clouds around T Tauri stars are revealed by infrared observations detecting thermal radiation from dust grains that have been heated by absorption of shorter wavelength radiation from the star. The nearest T Tauri stars are so far away (about 160 pc) that no direct observation of the inner parts of these clouds has been achieved, but some nearby young main sequence stars that are strong infrared sources, such as Vega (7.5 pc) and Beta Pictoris (17 pc), appear to be surrounded by molecular gas as well as dust. Somewhere like the Beta Pictoris disc (Plate 1.19) is an obvious place to look for the formation of planets.

ITQ 2.1 Optical images show that the outer edge of the Beta Pictoris circumstellar dust disc is about 220 AU from the star. Light from the star prevents the disc being seen closer in than about 66 AU, but infrared observations indicate that its inner edge lies about 20 AU from the star. How well does the suggested extent of this disc correspond with the distribution of planets and other bodies in our own Solar System?

Thus a circumstellar disc of dust like that around Beta Pictoris would extend inwards only to the orbits of the outermost planets in our Solar System, and fall well beyond the most massive ones, Jupiter and Saturn. The apparent absence of dust closer to Beta Pictoris than about 20 AU may be because dust that used to be in this inner region has already gone to form planets.

Although it seems likely that many main sequence stars of between about 1 and $3 M_{\odot}$ have planetary systems around them, we cannot obtain detailed views of planetary systems in the making, nor are there any *direct* observations of planets around any normal star other than the Sun. Thus our understanding of the genesis of planetary systems is based largely on deductions concerning the planetary system within which we live.

In this chapter we shall examine how the Solar System may have formed, about 4500 million years ago. We shall follow the story from the time of the circumstellar dust and gas cloud around the protoSun as far as the formation of the planets and their satellites. In Chapter 3 we shall begin to study what happens to a planet once it has formed.

We are using the conventional symbol M_{\odot} to indicate the mass of the Sun.

The age of the Solar System has been established mainly by radiometric dating of meteorites, which you will learn about in Chapter 8.

2.2 The solar nebula

2.2.1 Nebular theories

Direct observations of circumstellar discs, achieved with increasing clarity since the early 1980s, support a theory of planetary system formation that was proposed in its original form in 1796 by the French scientist Pierre Simon Laplace (1749–1827). Laplace suggested that the Solar System was formed by gravitational collapse of a large, initially spherical, rotating mass. According to this theory, the central region grew denser and became the Sun, and the remainder was forced into a disc of gas and dust, called the **solar nebula**, within

which the planets formed. This **nebular theory** has been with us ever since, and today receives more credence than other, more catastrophic, explanations such as the ripping out of solar material by tidal forces when another star passed close to the Sun.

However, even though most scientists believe that the Solar System formed from a nebula, there is no general agreement as to how it happened. There are many versions of the nebular theory, but it will not be possible to examine them here. Instead we will present you in this chapter with a story that fits most of the available evidence, and is at least moderately self-consistent. You should *not* regard it as representing a consensus view in all its details, and we will encourage you to be aware of difficulties and uncertainties when they crop up.

2.2.2 Physical processes in a collapsing cloud

Let us go back to the collapsing dust and gas cloud, of the sort that you encountered in Book 1, Subsection 3.2.3.

If this cloud were originally rotating a little, what would happen to the rate of rotation as the cloud became smaller but denser during collapse?

■ The cloud would have to spin faster, in order to conserve angular momentum (Book 1, Box 4.1).

Conservation of angular momentum can also explain the directions of rotation of the Sun, and the orbits of most planetary bodies in the Solar System. A remarkable aspect of the arrangement of the Solar System is that the orbits of all the planets, and most asteroids, lie close to the same plane (the ecliptic plane), and orbital motion and the Sun's spin are in the same direction (anticlockwise, as seen from above the Earth's North Pole). This direction of motion is called **prograde**, and is shared by the direction in which most planets rotate on their axes; the exceptions (see Table 1.1a) are Uranus and Pluto, which are tipped over on their sides, and Venus, which rotates very slowly backwards, or in a retrograde sense (expressed in the table by showing axial inclination as an angle greater than 90°).

The general sense of prograde motion close to a common plane would be a natural consequence of formation from a contracting, rotating cloud. The inward gravitational force on matter contracting within the plane of the cloud's rotation would be opposed by the rotation and by gas pressure, but the gravitational force on matter contracting from above either pole of the cloud would be opposed only by pressure, which would not become great enough to halt the polewards contraction until the nebula had become dense. Thus collapse would be easier near the poles, and the cloud would be flattened into a rotating disc, as shown in Figure 2.1 and Plate 2.1.

As you saw in Book 1, Section 3.2.3, the conversion of gravitational energy to thermal kinetic energy during contraction inevitably leads to heating. As the nebula became progressively denser in its centre, where what we may call the **protoSun** would be forming, it would become more opaque. This would inhibit heat-loss by radiation, and cause a substantial temperature rise. The pressure of this hot gas would eventually stop the gravitational contraction of the protoSun and of the innermost part of the disc. Gravitational contraction of the outer part of the disc would continue to be opposed mainly by the disc's rotation.



Figure 2.1 A cross-section through the solar nebula, at three stages in its contraction, as rotation forces it into a disc shape. The time difference between (a) and (c) was probably of the order of 10^5 years.

So far we have not considered the relative masses of the protoSun and the surrounding nebula. Here we touch on one of the more controversial aspects of Solar System formation. There are two major variants of the nebular theory. One is the low mass or **minimum solar nebula model**, which assumes that the mass of the nebula was equivalent to the present total mass of the planets with the addition of just enough hydrogen, helium and icy material to make the overall composition the same as the Sun today, a nebular mass of about $0.01 M_{\odot}$. On the other hand, the **massive solar nebula model** has the mass of the nebula around the protoSun to be of the order of $1.0 M_{\odot}$.

The minimum solar nebula model requires that all the suitable nebular material was converted into planets. In contrast, for the massive solar nebula model to be valid, we have to account for how a large proportion of its mass could have been lost. How this may have happened is tied in with the present distribution of angular momentum within the Solar System.

In Book 1, Subsection 4.4.1, you saw that when the core of a supernova collapses, conservation of angular momentum forces the resulting neutron star to spin very rapidly. Similarly, if the Sun is the product simply of contraction of the central part of a rotating nebula we would expect the Sun to be rotating rapidly. However, its present 25-day rotational period is far too slow, and although the Sun contains over 99.8% of the mass of the Solar System it possesses only a tiny fraction of the Solar System's angular momentum. Most of this resides in the orbital motions of the giant planets, as you can see in Figure 2.2.

You will see later in this chapter that hydrogen, helium and ice were prevented from being incorporated into planets with increasing efficacy the closer they were to the Sun.

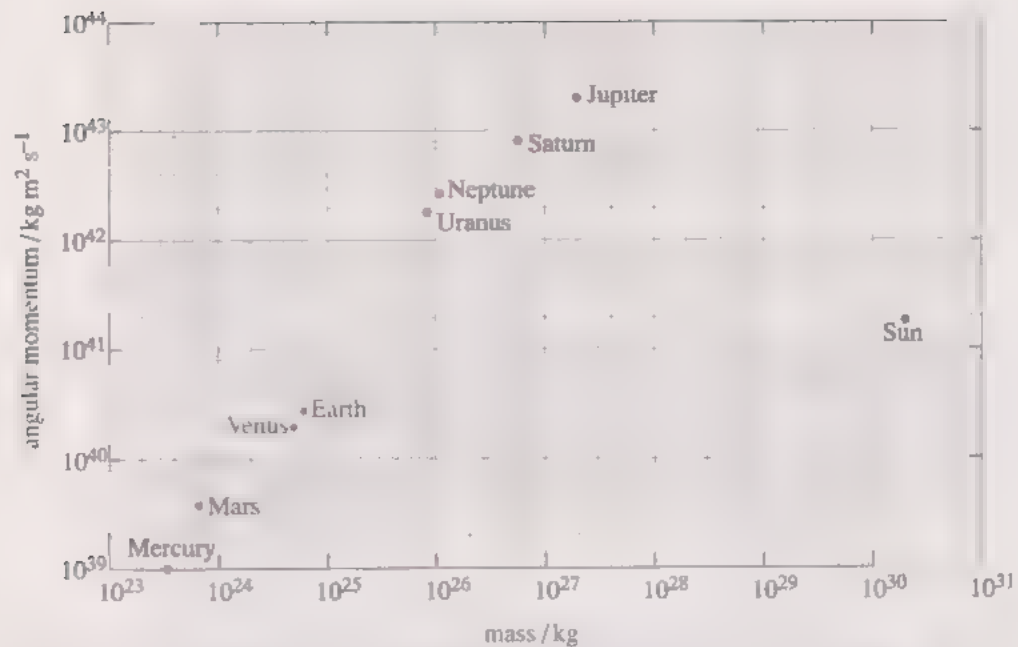


Figure 2.2 Angular momentum plotted against mass for the Sun (axial rotation) and the planets (orbital motion). Pluto's orbital angular momentum, and contributions to angular momentum by the axial rotations of the planets and the orbital motion of their satellites, are too small to show. Note that both axes are on a logarithmic scale. [Comment: The nearly straight line relationship for the planets between angular momentum and mass is because angular momentum is proportional to mass, but (as a consequence of Kepler's third law) it depends only weakly on orbital radius (in fact on its square root). If you very carefully draw a line at 45° passing through the data point for Mercury, you will find that the data points for the other planets fall progressively further from this line as orbital radius increases.]

For the Sun to have so little angular momentum, the matter that formed the protoSun must have transferred most of its angular momentum to the surrounding disc. This could not have happened if the gas molecules and dust grains in the solar nebula were moving independently in orbits dictated by Kepler's laws. They must have interacted, so that the disc as a whole behaved viscously; that is to say, there would have been **viscous drag** between adjacent elements of the disc. There are several factors that could have contributed to this. Probably the most important was **turbulence**, initiated by random motions inherited from the initial contraction.

ITQ 2.2 Consider a disc of gas and dust about the protoSun.

- What would be the relationship between the orbital periods of particles in circular orbits in the inner and outer parts of the disc, if Kepler's third law were applicable?
- What would tend to happen to the period of revolution in the inner and outer parts of the disc if these were linked by viscous drag, so particles could not orbit entirely independently?

Despite this effect, the whole of the disc could not be forced to rotate at a uniform rate unless it were an impossibly strong solid, so even allowing for viscous drag, the outer parts of the nebula would continue to rotate slower than inner parts. However, an important outcome of viscous drag would have been the outward transfer of angular momentum. Material in the outer part of the disc (having been speeded up) would flow outward, carrying angular momentum with it, but material in the inner part of the disc (having been slowed down) would fall inward onto the protoSun.

Thus, viscous drag within the solar nebula can explain the Sun's low angular momentum. The resulting outward flow at the outer edge of the nebula can account for the mass-loss that is necessary if we subscribe to a massive solar nebula model. There are two other ways in which the nebula could have lost mass at this stage: loss to space in bipolar outflow (Book 1, Subsection 3.2.3), and transfer of mass from the nebula to the protoSun, though neither explains the Sun's low angular momentum. There would have been further losses of both mass and angular momentum later, if the Sun went through a T Tauri phase of violent stellar winds (Book 1, Subsection 3.2.3).

It is much easier to explain the Sun's low angular momentum if large quantities of mass were lost from the solar nebula. This is a strong argument against the minimum solar nebula model, but we do not have to invoke the full massive solar nebula model in order to account for the present distribution of angular momentum. It seems reasonable to assume that the solar nebula began with a mass somewhere between that of the extreme models, say of the order of $0.1 M_{\odot}$.

2.2.3 *The composition of the solar nebula*

So far our discussion may have seemed rather hypothetical, divorced from observable and testable processes. We can get a fuller picture of events by examining evidence for the solar nebula's composition and how this may have varied with distance from the Sun.

So what was the solar nebula made of? Well, if we assume that the Sun was formed from the nebular material and accept that there has been no process operating that could have altered the abundances of the elements in the observable, outer, part of the Sun (Book 1, Chapters 1 and 3), then we can determine the solar nebula's composition, in terms of how much of each element it contained, by measuring elemental abundances in the outer part of the Sun.

□ How can we do that?

■ By spectroscopy, which you learned about in Book 1, Subsection 1.3.2.

But what evidence do we have that the solar nebula did have the same composition as the Sun? Well, the obvious way to test this is to compare the Sun's composition with something else that formed from the nebula, and see if they are similar. The easiest body in the Solar System to sample is the one we live on, the Earth. Unfortunately, as you will discover in Chapter 3, the rocks at the Earth's surface are unrepresentative of the Earth as a whole. However, a group of meteorites known as **carbonaceous chondrites**, about which you will learn more in Subsections 2.3.2 and 8.2.1, appear to come from bodies that have not experienced large-scale segregation of elements like that within the Earth. So if we want an independent check on the composition of the solar nebula, then carbonaceous chondrites are the most appropriate material to examine. Because they are thought to represent more-or-less unaltered samples of solar nebula material, carbonaceous chondrites are sometimes described as having **primitive** compositions. Figure 2.3 shows a comparison between the spectroscopically-determined abundances of elements in the Sun's atmosphere and the abundances of the same elements in carbonaceous chondrites, measured in the laboratory. Elements that form gases even at low temperatures (such as hydrogen and helium) are excluded from this plot because they are extremely rare within meteorites.

If you are unfamiliar with the chemical symbols for the elements, you will find these listed in the Appendix to Chapter 2 of Book 1.

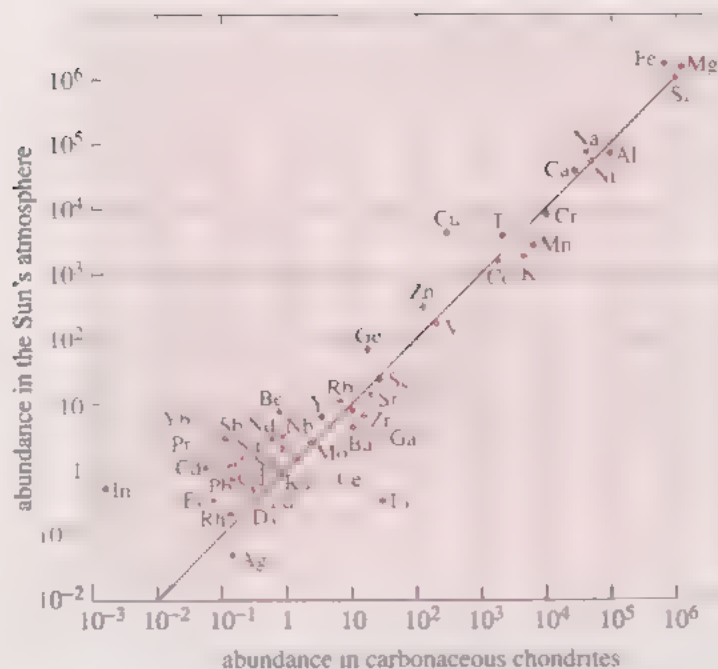


Figure 2.3 Comparison of the abundances of elements in the Sun's atmosphere and in carbonaceous chondrite meteorites. Both axes are on a logarithmic scale. By convention, the abundances are plotted as the number of atoms of each element for every 10^6 atoms of silicon. The straight line indicates where elements would plot if they had identical abundances in the Sun and in carbonaceous chondrites.

Figure 2.3 shows that there is clearly a very wide range of abundances of different elements in the Solar System.

How much less abundant is silver (Ag) than silicon (Si)?

The abundance of Si is (by convention) 10^6 , whereas the abundance of Ag is about 10^{-1} in both the Sun and in carbonaceous chondrites. Thus the abundance of Ag is about seven orders of magnitude less (i.e. seven powers of ten less) than the abundance of Si, which means there are about 10^7 times fewer atoms of Ag than of Si.

Despite the wide range of abundances, elements that are abundant in carbonaceous chondrites (e.g. Si, Mg, Fe) are also abundant in the Sun, and those that are rare in carbonaceous chondrites (e.g. Ag, Eu, Rh) are also among the rarest in the Sun. This can be seen by the fact that the elemental abundances fall fairly close to the straight line representing equal abundances in the Sun and carbonaceous chondrites. The relative abundances of the elements in the carbonaceous chondrites are often referred to as representing **chondritic composition**.

Thus, so far as we can tell by examining primitive meteorites, they have pretty much the same composition as the Sun (gas-forming elements excepted), and so appear to reflect the composition of at least that part of the solar nebula in which they formed. In the next section, we shall look at how this mixture of elements would have behaved as the nebula evolved.

Summary of Section 2.2 and SAQs

- 1 The available evidence is consistent with the Solar System having formed by contraction of a cloud of gas and dust known as the solar nebula. Rotation would have forced the nebula into a disc shape, and angular momentum was evidently transferred outwards from the protoSun into the nebula.
- 2 The least-altered (most primitive) meteorites, known as carbonaceous chondrites, contain elements (other than those which form gases) in similar abundances to the Sun. Elemental abundances in the Sun and carbonaceous chondrites probably reflect those in the solar nebula.

SAQ 2.1 (Objective 2.2) Why would a rotating solar nebula assume the shape of a disc, and what opposes the collapse of this disc?

SAQ 2.2 (Objectives 2.1 and 2.2) What processes in the solar nebula can account for the Sun's low angular momentum, and at what stage in the formation of the Solar System did each of these probably occur?

2.3 Condensation and coagulation

Until now, we have considered the composition of the solar nebula only in terms of abundances of elements. At temperatures below those of stellar atmospheres, most elements will form chemical compounds, including some well known minerals. The behaviour of these compounds will have controlled the evolution of the solar nebula, and some important examples are given in Box 2.1.

Box 2.1 Minerals and chemical compounds

As described in *Preparatory science*, Subsection 2.2.3, atoms of most elements form chemical compounds. A compound has a specific formula, involving simple multiples of its constituent elements. Compounds that are important in the evolution of the solar nebula include water (H_2O), methane (CH_4), troilite or iron sulphide (FeS), corundum (Al_2O_3), and pyroxene ($\text{CaMgSi}_2\text{O}_6$).

Naturally-occurring compounds that are found within rock are referred to **minerals**. Many ionic substances (*Preparatory science*, Section 2.5) that are minerals do not obey a precise formula. For example, a crystal of pyroxene ($\text{CaMgSi}_2\text{O}_6$) can have any proportion of iron (Fe) ions substituted in place of magnesium (Mg), and both iron and magnesium can substitute for calcium (Ca). All these ions have a charge of +2, so their relative abundances do not affect the balance of positive and negative ions. Moreover they are sufficiently similar in size that their substitution does not affect the atomic-scale structure of the crystal. The mineral pyroxene can therefore be represented by the general formula $\text{X}_2\text{Si}_2\text{O}_6$, where X represents an atom of either Ca, Fe or Mg, so long as there is a total of 2 atoms of X for every 2 atoms of silicon (Si) and 6 of oxygen (O). A mineral having Si and O in its formula, which is typical of most common minerals in rock, is called a **silicate**.

So in what forms were the elements in the solar nebula? Were they present as gases consisting of atoms and simple molecules, or did they constitute tiny solid particles? The answers to these questions depend mainly on the temperature within the nebula, which must have been greatest near the centre, and will have changed over time. Astronomical observations are much better at determining the sizes of grains in circumstellar dust clouds (typically 1–30 μm around Beta Pictoris and T Tauri stars), than their compositions (see, for example, Book 1, Subsection 5.2.5), and unfortunately it is not clear whether such grains are dominated by ices, carbon-rich material or silicate minerals. Let us now pursue the story of the solar nebula, by considering how solid material might have begun to be gathered together.

2.3.1 Grains in the solar nebula

The solar nebula must have grown hotter as the nebula became too dense for radiation from the protoSun to escape directly to space. This heating would have been augmented by viscous drag, and continued conversion of gravitational energy to thermal kinetic energy. Evidence from meteorites suggests that the temperature never rose above about 400 K at about 4–5 AU from the protoSun. However it is likely that the maximum temperature about 1 AU from the protoSun was around 2000 K, near the time of stage (b) in Figure 2.1. This would have been sufficient to break any pre-existing grains (of virtually all feasible compositions) into their constituent atoms. At this stage, most of the elements within 2–3 AU of the Sun would have been in the gaseous state, as dissociated atoms, though there would have been some simple molecules such as H_2 and CO. Subsequently, as the nebula became more transparent to solar radiation and as the other heat sources ceased to act, it would have begun to cool. This would have allowed new compounds to form, growing as dust grains, by **condensation** (Box 2.2).

Box 2.2 Two types of condensation

You probably know condensation as what happens when you breathe out on a cold day, and water vapour condenses to form the tiny droplets that make your breath visible. On cooling, molecules of H_2O in the gaseous state lose energy and come together to form a liquid. This type of condensation is characteristic of molecular substances (*Preparatory science*, Section 2.5). However, under low pressure – as in the solar nebula – gases of molecular substances condense directly to form solids, with no intervening liquid phase.

Condensation of ionic substances is a different process and usually occurs at higher temperatures (*Preparatory science*, Section 2.5). In the solar nebula, elements that had existed as free atoms in the nebular gas at high temperature would have joined together as the nebula cooled to form chemical compounds, involving two or more elements. Once a tiny grain has condensed in this way, it is easier for individual free atoms to attach to its surface, rather than to start a new grain.

The nature of the starting material from which the planets in the inner Solar System developed was probably dictated by the highest temperatures reached locally. Further out from the protoSun it becomes more likely that pre-existing grains could have survived, and you will examine some of the evidence for these in certain meteorites in Section 8.3.

The sequence in which solid particles would have condensed while the solar nebula was cooling depends on a complex interplay of the *stability* of compounds and the *rate* at which they could form from a gas at the prevailing temperature.

and pressure. A simplified **condensation sequence** is suggested in Table 2.1, showing the temperatures at which some notable constituents of the Solar System probably began to condense.

Minerals and other compounds that condense at high temperatures are described as **refractory**, and those that do so at low temperatures are described as **volatile**. The condensation sequence is the reverse of the evaporation sequence, that is to say if the temperature were to increase, the most volatile compounds would volatilize (form gases) first whereas the more refractory compounds would be the last to do so.

There is no hard-and-fast rule quantifying what is meant by 'high' or 'low' temperature; it depends on context: a substance that is described as volatile in one context may legitimately be called refractory in another.

Table 2.1 Simplified condensation sequence for notable substances forming from the solar nebula. The most refractory (least volatile) substances are at the top, becoming less refractory (more volatile) towards the bottom. The temperatures shown are those at which substances would condense in the solar nebula if the pressure were 10^{-3} bar (10^2 Pa). At lower pressures, condensation temperatures would be slightly lower.

Temperature / K	Substance	
1 758	corundum, Al_2O_3	
1 647	perovskite, CaTiO_3	
1 513	spinel, MgAl_2O_4	
1 471	nickel-iron metal, Ni,Fe	
1 450	pyroxene (diopside), $\text{CaMgSi}_2\text{O}_6$	
1 444	olivine (forsterite), Mg_2SiO_4	
< 1 000	alkali feldspars, $(\text{Na,K})\text{AlSi}_3\text{O}_8$	
700	troilite, FeS	
550–330	hydrated minerals ^a	ionic substances↑
180	water, H_2O (as an ice)	molecular substances↓
120	ammonia, as $\text{NH}_3\cdot\text{H}_2\text{O}$ (ice)	
70	methane, as $\text{CH}_4\cdot 6\text{H}_2\text{O}$ (ice)	
70	nitrogen, as $\text{N}_2\cdot 6\text{H}_2\text{O}$ (ice)	

^a Hydrated minerals are chiefly silicates with OH or H_2O in their formulae.

It is generally believed that the planet-forming process began with the cooling of the solar nebula after it had reached its peak temperature. We have to imagine progressively more dust-sized grains forming by condensation within the nebula at about stage (c) of Figure 2.1. As these grains formed, they would have settled towards the mid-plane of the nebula, producing a sheet rich in dust and larger particles. This settling would be possible only after turbulence within the nebula had declined, and it would have become more rapid as the grains grew. It is presumed that the discs around Beta Pictoris and other young stars are the remnants of such sheets.

2.3.2 Coagulation of grains

Now we have to consider how these grains could continue to grow, especially after much of the nebular material had condensed. Gravitational attraction between individual dust grains is too slight to bring grains together. Instead, chance encounters seem to have been the most important factor. In a collision between two particles, any of three things can happen: one or both particles can be fragmented (which is the opposite of what we are looking for), they can bounce off each other (which would not help us), or they can stick together

An alternative term for the coagulation process is cohesion.

In your general reading, you may find column mass called column density. The two terms mean the same.

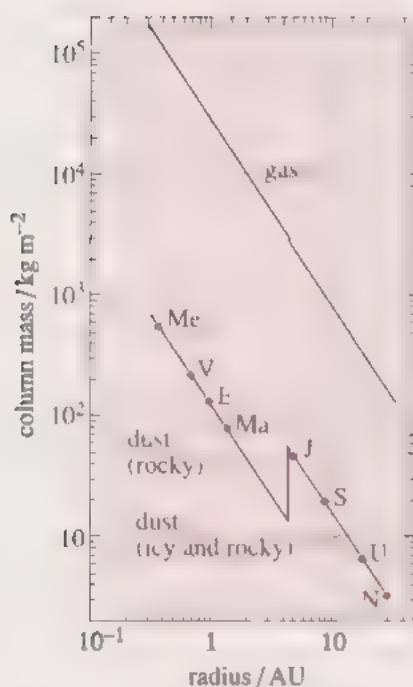


Figure 2.4 The column mass of gas and dust in the solar nebula (at a stage represented by Figure 2.1c) plotted against distance from the protoSun. Obvious abbreviations are used to indicate the positions of planets that ultimately formed. The reason for the increase in column mass of dust near 5 AU is suggested in the text. (Smaller steps associated with the condensation of more-volatile compounds such as ammonia have been omitted.)

(which is just what we want). Fragmentation would be common only if the collisions were more violent than the sort of chance encounters likely in the solar nebula. One suggestion as to why colliding grains stuck together instead of bouncing apart is that their surfaces were fluffy, perhaps riddled with cavities caused by evaporation of ice that had previously condensed there. Magnetism and small electric charges may also have helped to hold grains together. Whatever the mechanism, this sticking together of grains that had grown by condensation in the solar nebula is generally referred to as **coagulation**.

The rate at which coagulation could proceed must have depended on the density of material within the nebula. One model suggests that once settling of dust grains towards the mid-plane had begun, it would have taken only about 2 000 years to produce particles up to 10 mm in diameter at 1 AU from the Sun, about 5 000 years to produce particles up to 15 mm in diameter at about 5 AU (the Jupiter region) and about 50 000 years to produce particles up to 0.3 mm in diameter at about 30 AU (the Neptune region).

There are three factors that would contribute to the slower growth of grains further from the protoSun. First, **column mass** decreases outwards, so grains might be expected to condense more slowly in the outer reaches of the nebula (in this context, **column mass** is the mass per square metre of the solar nebula, measured perpendicular to its plane). Second, the lower density would result in less frequent collisions and hence slower coagulation. Third, the nebula was probably fatter in its outer reaches (Figure 2.1b–c), and so the majority of grains forming there would have to fall further to reach the mid-plane, taking longer to **complete their journeys**. Estimates of the rate of coagulation depend on assumptions made about the densities of both gas and dust present in the solar nebula. A reasonable model is shown in Figure 2.4, which shows the variation in column mass against distance from the centre. Note that this model has several orders of magnitude more mass present in the form of gas than as dust, especially in the inner Solar System.

Given that this figure represents the situation after everything able to condense from the nebula has already done so, can you suggest what remains in the gas?

■ The gas will be made up of the abundant volatile substances, notably hydrogen (as H_2 molecules) and atomic helium (He), neither of which would have condensed.

The step up in the column mass of dust near 5 AU is anticipated because at the relevant time the temperature at this distance would have been low enough for water to condense, in the form of ice. Note that there is a downward step in the column mass of the gas at the same distance that (allowing for the logarithmic scale) exactly compensates for the increase in column mass of dust.

Estimating the rates of processes such as condensation within the solar nebula is fraught with difficulties, but the icy nature of most of the major satellites of the outer planets demonstrates that water did condense in considerable quantities near Jupiter and beyond.

□ Can you suggest how it is possible for water to have been incorporated in material that condensed at this stage considerably *closer* to the Sun than 5 AU?

■ It can be incorporated in hydrated minerals, which condense at higher temperatures than that at which water condenses to ice (Table 2.1). (This is probably the origin of much of the water in the Earth.)

A possible window back to the time when coagulation had led to particles a few millimetres across is provided by what may be relicts of these particles embedded within certain meteorites. There is a bewildering variety of meteorite types, for

which you will get some feel in Chapter 8. For now we will just note those that consist of globules of silicate minerals a few millimetres across embedded in a matrix (also dominantly of silicates). Such a globule is called a **chondrule**, and the meteorites in which they are found (Plates 2.47 and 2.48) are known as **chondrites**, or chondritic meteorites. Carbonaceous chondrites (Subsection 2.2.3) are the least-altered (most primitive) variety. For now the important thing is that chondrules *could* represent coagulated grains, although we cannot be sure. We will consider how chondrules and matrix could have been gathered together in the next section.

We have now traced the story of the early Solar System from the origin of the solar nebula up to the formation of particles a few millimetres in size. The time required for this appears to have been surprisingly brief. Condensation probably took just a few thousand years, and you will see in Chapter 8 that meteorites provide isotopic evidence that this must have occurred within a few million years of the creation of the nebular material. We have not considered in any detail whether or not the processes we have discussed are likely to have been going on simultaneously in all parts of the solar nebula, but you have seen examples of both time-dependent and space-dependent variations. For example, in the inner Solar System the temperature was always too high for water to condense as ice. Also, things will have happened more slowly in the outer Solar System for two reasons: the lower temperature would decrease the speed of chemical reactions, and the lower density of the nebula would give grains less frequent opportunities to coagulate. In the next section we will consider how the grains could come together to form the bigger bodies necessary to create the planets.

In less-primitive chondrites, new minerals have replaced the original ones, showing that both the chondrules and the matrix were heated after they were assembled.

Summary of Section 2.3 and SAQs

- 1 Much of the original dust in the solar nebula may have been vaporized, especially in the inner Solar System, principally as a result of heat released by loss of gravitational energy within the nebula and within the protoSun. Nebular material would have begun to condense as the temperature dropped, with the refractory substances appearing first and the more volatile substances condensing later, according to the condensation sequence. Within about 5 AU of the Sun, although some water will have become trapped within hydrated minerals that condensed, the temperature never dropped low enough for ice to condense.
- 2 Grains of dust that condensed from the solar nebula settled towards the central plane of the disc. When they came into contact they tended to do so gently, and stuck together (coagulated) to form larger particles. This sticking together may have been encouraged by fluffy surface textures, magnetism and electric charges.

SAQ 2.3 (Objectives 2.1 and 2.2) Many chondrules contain small amounts of iron sulphide (FeS) mixed in with the silicate minerals. If we assume that the iron sulphide has not grown within the chondrules subsequently, what constraint does this put on the temperature at which the chondrules formed?

SAQ 2.4 (Objectives 2.1 and 2.2) According to the model in Figure 2.4, what are:

- (a) the total proportion of rock and ice expressed as a percentage of the total material in the solar nebula?
- (b) the ratio between rock-forming and ice-forming material in the nebula?

2.4 Planetesimal formation

As you learned in Book 1, Subsection 3.2.3, young stars of around $1 M_{\odot}$ experience a T Tauri phase of violent stellar winds, probably when they are about a million years old. As you will see in Subsection 2.5.3, there is circumstantial evidence that the Sun itself went through such a phase. Any body less than about **10 m across would have been blown out of the Solar System by this T Tauri wind**, so the material destined to form the planets must have been gathered into masses of at least that size by then. Indeed it is widely believed that the next stage in planetary formation after initial coagulation led to the growth of bodies about 0.1–10 km across, known as **planetesimals**. One model for how this happened calls for a continuation of coagulation during collisions in the central dust sheet. Another theory is that once this sheet had contracted to less than about 100 km thick it became gravitationally unstable, breaking into a multitude of turbulent knots, each of which developed into a planetesimal. Either way, it appears that coagulated grains of all sizes were gathered together to form planetesimals, probably within a few hundred thousand years after condensation began.

Can you suggest what could become of any dust trapped between larger coagulated grains while a planetesimal was forming?

■ This material could become the matrix between chondrules in a chondritic meteorite.

The complex histories suggested by the mineralogy and texture of most meteorites suggest that they are fragments of planetesimals or larger bodies that were broken apart by collision – but the least-altered varieties are the best relics available of the process of planetesimal formation.

This Section is summarized at the end of Section 2.5.

2.5 Planetary embryos

2.5.1 Runaway growth

Once a planetesimal has reached about 10 km across it has become massive enough that its own gravitational attraction is sufficient to perturb the motions of other planetesimals. A larger planetesimal will deflect the trajectories of smaller planetesimals towards it (a phenomenon known as **gravitational focusing**) resulting in increasingly frequent collisions. Provided most of the material in two colliding planetesimals ends up together, a process known as **accretion**, instead of fragmenting and dispersing, the bodies that remain will get progressively larger, and the large bodies will become fewer in number.

It is impossible to model exactly how planetesimals would have interacted, because there were far too many of them and there are factors of unknown magnitude, such as the degree of viscous drag due to the remaining gas, that should be taken into account. However one important outcome is clear: because of their greater strength of gravitational focusing, larger planetesimals would have grown more rapidly than smaller ones, while the total mass incorporated in smaller bodies declined. It is probable that a situation of **runaway growth** developed in each neighbourhood, in which the growth of one planetesimal outpaced that of all its rivals, allowing it become a much larger object known as a **planetary embryo**.

It would not be possible for colliding planetesimals to stick together *without* fragmenting, because the collisions would be too violent. (By contrast, grains *can* stick without fragmenting – see Subsection 2.3.2.) After a collision therefore, the new, larger, planetesimal is made of a mixture of accreted fragments from the two planetesimals that collided.

ITQ 2.3 Assume the density of the material forming planetesimals was $3 \times 10^3 \text{ kg m}^{-3}$. How many 10 km diameter planetesimals would it take to form the Earth, which has a mass of $6.0 \times 10^{24} \text{ kg}$? [Comment: It is a reasonable approximation to treat these planetesimals as spheres. The volume V of a sphere of radius R is given by $V = \frac{4}{3}\pi R^3$.]

2.5.2 Planetary growth in the inner Solar System

The style and speed of runaway growth is likely to have varied with distance from the Sun, slowing drastically in each vicinity as the supply of smaller bodies available for capture became depleted. Limiting masses calculated for planetary embryos at different distances from the Sun are indicated in Figure 2.5.

We will follow events in the inner Solar System first, before considering what happened further out. In the region where the terrestrial planets are now found (from Mercury to Mars) one planetary embryo probably grew about every 0.02 AU outward from the Sun. One of the many models describing the evolution of a swarm of planetesimals into a planetary embryo in this region is illustrated in Figure 2.6. Please study this figure and its caption; it is probably correct in a general sense, though its details are not to be trusted.

If the models illustrated in Figures 2.5 and 2.6 are even roughly correct, runaway growth appears able to have produced substantial planetary embryos before the Sun's T Tauri stage. However the supply of planetesimals within range of each embryo's gravitational influence was sufficient only to allow it to reach about a tenth of the mass of the individual planets that would eventually form.

After this, planetary growth can have proceeded only at a slower pace, as a result of chance collisions between planetary embryos whose orbits crossed. Whenever two embryos collided, the smaller of the two would have been fragmented and mostly accreted onto the larger one. A collision between two embryos of similar mass would probably have broken both into a swarm of debris, most of which would have rapidly come together (accreted) under its own gravitational attraction. Either way, the result would be the same, fewer and fewer remaining embryos that would become more massive after each collision. Calculations for Mercury, Venus, the Earth and Mars suggest that in each case it

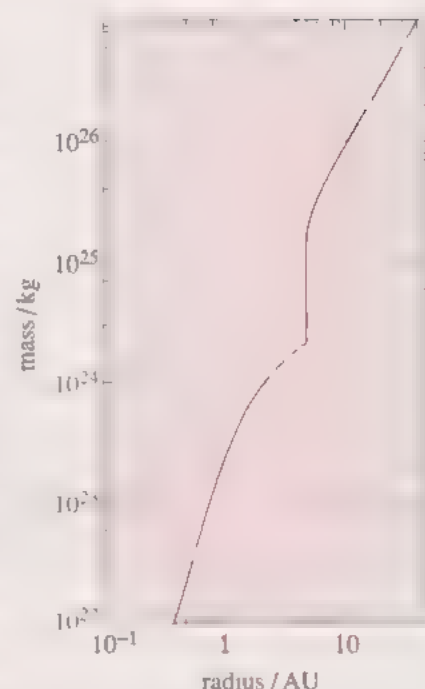


Figure 2.5 The final masses for the largest planetary embryos produced by runaway growth at different distances from the Sun. The step-up at about 5 AU is because of the condensation of water-ice beyond that distance (see Figure 2.4).

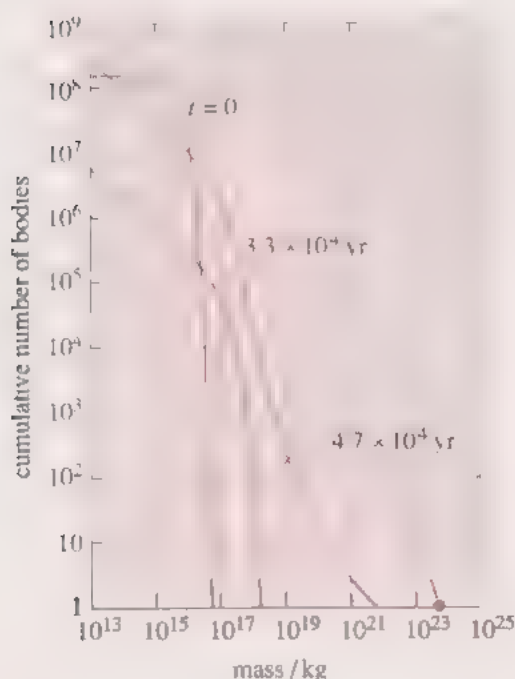


Figure 2.6 The evolution of a swarm of planetesimals orbiting the Sun between 0.99 and 1.01 AU. The curves show the distribution of planetesimals by mass at three different times, beginning at $t = 0$ with almost all the planetesimals between about 10^{16} kg and 10^{17} kg in mass (about 10 km in diameter). At 33 000 years, the largest planetesimal is just under 10^{22} kg , and represents the high-mass end of a continuum of planetesimal masses. However, after 47 000 years runaway growth has led to the production of a single planetary embryo between 10^{23} and 10^{24} kg in mass. By that stage, no other body in this region of the Solar System has a mass within five orders of magnitude of this. Note that the axes are logarithmic, and that the vertical axis shows the cumulative number of bodies greater than the mass plotted on the horizontal axis.

would take about 10^7 years for growth to about half the final planetary mass and about 10^8 years for growth to be essentially complete (Figure 2.7). We will examine the curious case of the Moon in Section 3.4.

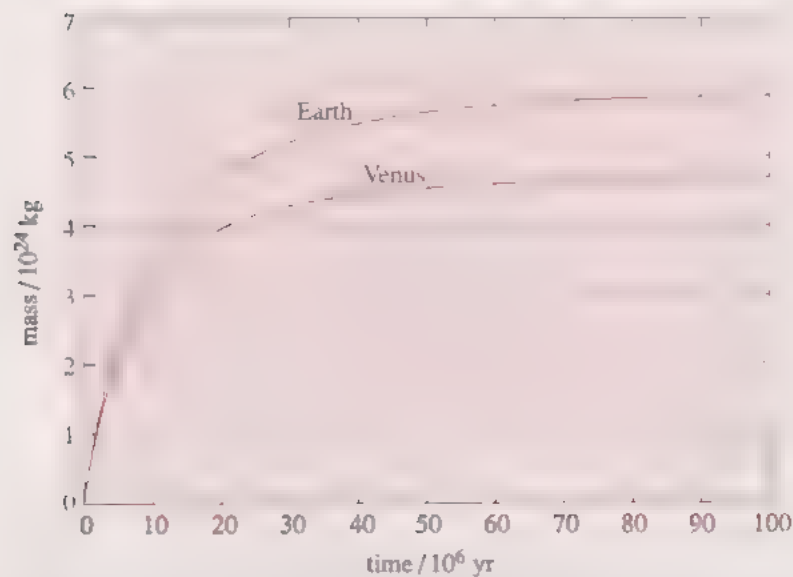


Figure 2.7 A model for the timescale for the growth of the Earth and Venus as a result of collision between planetary embryos. Mercury and Mars probably grew at similar rates.

2.5.3 Formation of the giant planets

Turning now to the outer Solar System, as you can see from Figure 2.5 the expected masses of planetary embryos in the Jupiter–Neptune region (5–30 AU) are at least two orders of magnitude larger than those at 1 AU. However, the large volume of space occupied by the planet forming material in the outer Solar System makes it more difficult for runaway growth to begin (indeed, some authorities dispute that it could have happened at all), and when it did get underway it would produce fewer but larger planetary embryos. Figure 2.8 shows a simulation of how runaway growth might proceed at 5 AU.

According to this, an embryo (probably a mixture of rocky and icy materials) with a mass somewhat less than 10^{26} kg, about $10 M_{\oplus}$, is produced after 3.9×10^5 years. (This is about ten times as long as it would take to produce a planetary embryo at 1 AU, according to the comparable simulation shown in Figure 2.6.) This hypothetical $10 M_{\oplus}$ body would act as a **kernel** that would, with the aid of gravitational focusing, sweep up planetesimals, smaller debris and nebular gas, leading to the formation of the planet Jupiter. As this body grew, it would be able to capture a significant amount of gas from the solar nebula, and its gravitational attraction would be adequate to collapse the captured gas inwards, producing an envelope consisting largely of molecular hydrogen (H_2) and atomic helium (He).

Runaway growth such as this would take even longer to produce the kernels for Saturn, Uranus and Neptune, probably about 2, 10 and 30 million years respectively. This increasing timescale with distance from the Sun is probably the key to understanding the outward trend in composition among the giant planets.

We are using the conventional symbol M_{\oplus} to indicate the mass of the Earth.

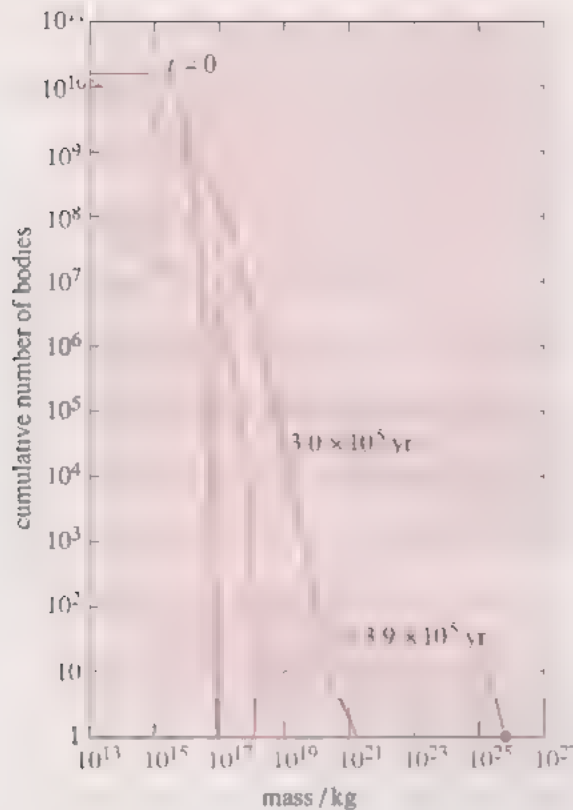


Figure 2.8 The evolution of a swarm of planetesimals near 5 AU leading to the production of a Jupiter kernel by runaway growth. Compare with Figure 2.6.

Look at the data for Jupiter, Saturn, Uranus and Neptune in Table 1.1a. Can you see any trend in their *densities* as their distance from the Sun increases?

■ No, there isn't one.

The reason is that the more massive a planet the stronger its gravity and the more compressed, and therefore denser, will be the material in its interior. If this **self-compression** is taken into account, then the densities of the four giant planets can be explained by their containing roughly similar amounts of rock and ice ($10\text{--}20 M_{\oplus}$) but diminishing amounts of hydrogen and helium further from the Sun. Jupiter is thought to contain about $200 M_{\oplus}$ of H_2 and He, Saturn about $70 M_{\oplus}$, and Uranus and Neptune about $1 M_{\oplus}$ each.

Can you suggest an event that would have prevented the more distant planets from gathering as much gas from the solar nebula as Jupiter did?

■ If their kernels formed later (as we have suggested) then maybe most of the gas in their neighbourhood had been removed before they grew massive enough to attract much gas gravitationally. The most likely explanation for this removal is the T Tauri wind.

This is the circumstantial evidence that the Sun did pass through a T Tauri phase that we referred to in Section 2.4. The T Tauri wind blew away the remaining H_2 and He, so that no more gas could be captured by the giant planets. A related argument is that the Earth should have captured an atmosphere of about $0.03 M_{\oplus}$ from the solar nebula. This is 3×10^5 times the mass of its present atmosphere, which furthermore has a totally different composition. We can conclude that the Earth's primitive atmosphere (and those of the other terrestrial planets) was probably lost during the T Tauri phase. You will study the processes responsible in Subsection 6.6.2.

2.5.4 Origin of the asteroids and comets

Jupiter is much the most massive body in the Solar System after the Sun, and its gravitational field is likely to have been an influential factor in the history of the Solar System ever since the development of the massive Jupiter kernel postulated in the previous subsection. The most obvious sign of Jupiter's influence can be seen in the absence of a substantial planet between Mars and Jupiter, even though simulations of runaway growth suggest the appearance there of planetary embryos in excess of 10^{24} kg within 4×10^5 years. Today this region is populated only by the asteroids, all less than about 10^{21} kg in mass (see Table 1.1c), and amounting to a total mass of no more than about 3×10^{27} kg (about $5 \times 10^{-4} M_{\oplus}$). It was once thought that the asteroids are the wreckage of a planet that was destroyed by a collision. Now it seems more likely that the gravitational influence of the growing Jupiter kernel perturbed the orbits of planetesimals orbiting near 3 AU, increasing their eccentricities. This would have ensured that when these planetesimals collided with each other the velocity, and hence the kinetic energy, of the impact was so great that fragmentation and dispersal became just as likely as growth by accretion.

As a result, runaway growth never took place, and the orbital perturbations probably led to much of the original material here being scattered. By now most of this debris has impacted onto other bodies, forming the impact craters that characterize old surfaces throughout the Solar System (Chapter 4). The irregular shapes of the surviving asteroids – which are essentially those planetesimals to have survived – attest to their violent pasts (Figure 2.9, Plate 2.43).

Jupiter may also have played a role in the origin of comets. These may originally have been icy planetesimals that were flung outwards as a result of a close encounter with Jupiter or one of the other giant planets (see Figure 2.10). Such an encounter could fling a planetesimal off in a very elongated orbit extending out to the region of the Oort cloud. A rival explanation is that comets could have condensed in the outermost reaches of the Solar System in the first place. Evidence for growth of bodies beyond the orbit of Neptune was provided by the discoveries in August 1992 and March 1993 of two bodies (1992 QB₁, named Smiley, and 1993 FW, named Karla) in near-circular orbits at about 44 AU and each estimated to be about 200 km in diameter. Four smaller ones, orbiting at about 34 AU, were discovered in September 1993. In Section 8.6 you will have a chance to see what light can be shed on the origin of cometary material by our knowledge of its chemical composition.



Figure 2.9 Views of the Earth-crossing asteroid Toutatis obtained by radar as it passed about 4 million kilometres from the Earth in December 1992. It appears to consist of two irregular objects (about 4 km and 5 km across) that are joined together, and rotates in a period of about 10 days. Each part is cratered as a result of previous collisions.

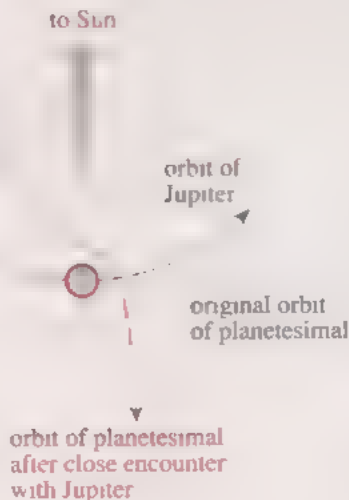


Figure 2.10 A possible explanation for the origin of comets. A close encounter with Jupiter perturbs the orbit of an icy planetesimal, sending it out towards the Oort cloud.

Summary of Sections 2.4 and 2.5 and SAQs

- 1 Within a few hundred thousand years of the creation of the solar nebula, grains in the dust disc that had formed near its central plane became gathered into masses about 0.1–10 km in size known as planetesimals, either because of gravitationally-driven turbulence or as a result of continued coagulation.
- 2 Once a planetesimal had reached about 10 km in size, its gravitational field would have become sufficient to encourage collisions, by means of gravitational focusing. Eventually, the growth of one body in a vicinity would have outpaced all the others (runaway growth), producing a planetary embryo. At about 1 AU these were probably about a tenth of the mass of the Earth. The creation of planetary embryos probably took of the order of 10^4 or 10^5 years in the inner Solar System, but the timescale goes up by roughly an order of magnitude in turn for each of the giant planets.
- 3 Each terrestrial planet grew by a series of collisions between planetary embryos that took about 10^8 years to complete. Runaway growth took longer to get established in the outer Solar System, but the planetary embryos that grew there became sufficiently large that they could act as kernels around which large masses of gas were captured directly from the solar nebula.
- 4 The onset of a powerful solar wind as the Sun entered its T Tauri phase swept the remains of the solar nebula outwards into space, and this evidently happened early enough to forestall Saturn, Uranus and Neptune from capturing as much gas as the earlier-formed Jupiter was able to. Loss of the primitive atmospheres of the terrestrial planets can be attributed to the T Tauri wind, but the giant planets were able to hold onto most of the gas they had already captured.
- 5 The sub-planetary size of the asteroids can be attributed to gravitational perturbation of the orbits of planetesimals in this region by the incipient Jupiter, which ensured that mutual collisions were too energetic to lead to runaway growth. Comets are icy bodies that may have formed in the outer reaches of the Solar System, or they could have been flung outwards after passing close to Jupiter or another giant planet.

SAQ 2.5 (Objective 2.2) The simulations in Figures 2.6 and 2.8 both begin with a population of planetesimals between about 10^{-6} and 10^{17} kg in mass (i.e. about 10 km diameter). The flat tops to the curves for $t = 0$ show that there are no planetesimals less than about 10^{16} kg in mass. In both simulations, at times later than $t = 0$ the low-mass ends of the curves slope down to the right. What does this tell you about the masses of the smallest bodies present at these later times, and to what process do you attribute this?

SAQ 2.6 (Objective 2.2) What factors are responsible for the giant planets being richer in volatiles such as hydrogen and helium than the terrestrial planets?

2.6 Satellite systems

We have now described processes that can account for the origin of all bodies in the Solar System except planetary satellites. We will leave the Moon aside until Section 3.4, and concentrate on the major satellites of the giant planets. With the exception of Neptune's large satellite Triton, these are in orbits close to their

planet's equatorial plane, travelling in the same direction as the planet rotates. It is unlikely that these satellites could have formed elsewhere and been captured later by their planets, because in that case their orbits would not be so well-ordered. Instead, the satellites probably condensed from a disc of gas and dust about each primitive planet, rather like the solar nebula in miniature. It is debatable whether a such a **protosatellite disc** is more likely to have been shed by a young planet as it contracted, or to have been gathered around it by scavenging of stray planetesimals and other material from the remains of the solar nebula.

Whatever the origin of a protosatellite disc, we can imagine it developing into individual satellites in much the same way as the planets had formed. Small particles would have aggregated into larger ones that would, over time, have collided. Eventually the largest bodies would have swept up all the debris in their vicinity to form the surviving satellites. The protosatellite disc would have been rotating in the same direction as the solar nebula, and this explains the satellites' consistent orbital directions.

We can learn more about the origin of planetary satellites by examining the four large satellites of Jupiter, known as the **galilean satellites**, after Galileo who discovered them in 1610 using one of the first telescopes. To see what these tell us, you should now do ITQ 2.4.

ITQ 2.4 (a) Data on the galilean satellites (Io, Europa, Ganymede and Callisto) are included in Table 1.1b. Can you see any systematic trend in their densities?

(b) The density of ice is about $1.0 \times 10^3 \text{ kg m}^{-3}$, whereas that of any rock within these bodies is likely to be about $3.0 \times 10^3 \text{ kg m}^{-3}$. What does this extra information tell you about the probable compositions of each of these satellites?

Spectroscopic studies of sunlight reflected from the surfaces of Ganymede (Figure 2.11) and Callisto demonstrate that their surfaces at least are dominated by water-ice. Similar data show that ice is prevalent over the whole of Europa too (Plate 2.22), though Europa's high density demonstrates that its coating of ice can be no more than a few tens of kilometres thick, and its interior is probably mostly rock. Io is an essentially rocky body with no surface ice (Plates 2.26–2.28). Its high density shows that its interior contains even denser material, probably in the form of an iron-rich core.

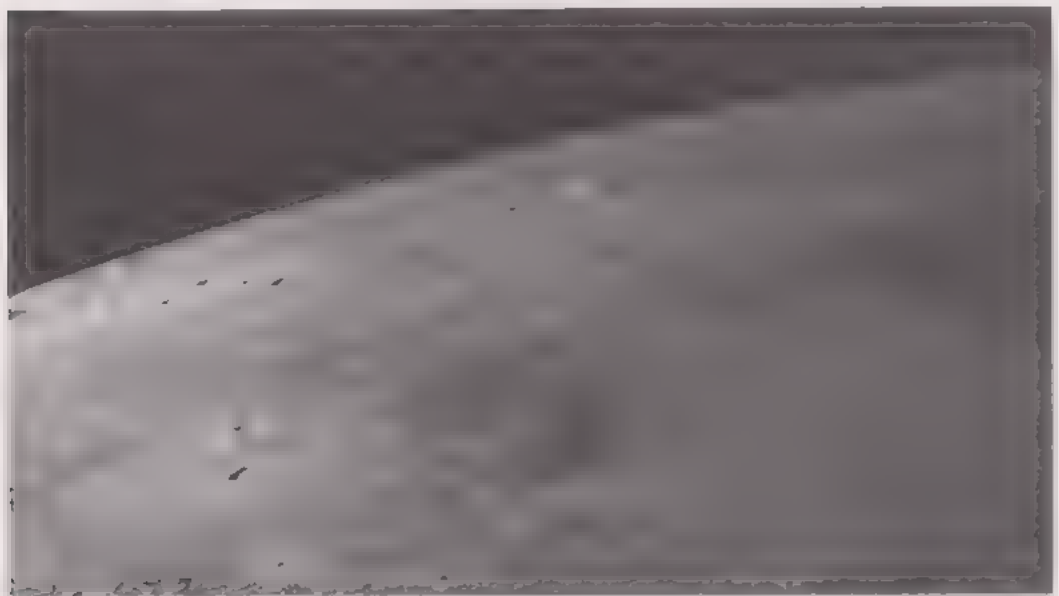


Figure 2.11 A close-up view of Ganymede. Spectroscopic studies show that the surface is icy, despite the craters and other features that would seem to suggest a rocky surface.

Thus the galilean satellites resemble the Solar System in miniature. The inner ones are less massive and denser, whereas the outer ones are more massive but less dense.

Can you suggest a condition within the protosatellite disc around Jupiter that could account for this density trend?

■ Perhaps it was too hot in the inner part of this disc for ice to exist (i.e. above about 180 K, according to Table 2.1).

It is now accepted that heat generated by the conversion of gravitational energy to thermal kinetic energy as material in the protosatellite disc fell in towards the primitive Jupiter probably raised the temperature in the inner part of the disc. This either prevented ice from condensing there, or vaporized any ice that had already condensed. As you will discover in Chapter 7, even today Jupiter radiates more energy than it receives from the Sun.

There is no such density trend within the satellite systems of Saturn and Uranus. All their major satellites consist of a rock plus ice mixture, and none are essentially rocky like Io and Europa. However, this cannot be used as evidence that they did not form in a protosatellite disc, because – being less massive, the primitive Saturn and Uranus would have generated less heat in infalling material than Jupiter. **An additional factor that could be responsible for the lack of consistent density trends is that most of these satellites are much smaller than the galilean satellites, so the inner ones in particular would have been more vulnerable to fragmentation by impacts involving large comets or other bodies.** After fragmentation, but before re-accretion to form the satellites we see today, material from different satellites could become mixed. Although they are smaller, the satellites of Saturn and Uranus are similar in composition to Ganymede and Callisto, except that, having formed further from the Sun and consequently in a lower temperature environment, they were probably able to incorporate methane and ammonia into their ice (see Table 2.1).

11Q 2.5 Can you think why inner satellites are more vulnerable to impacts than outer ones, and why smaller ones are more likely to break up as a result of an impact?

Satellites less than about 200 km in radius are mostly irregular bodies (e.g. Figure 2.12), and may be fragments resulting from such collisions. Collisions may also be the origin of the rings of debris that are now known to encircle each of the giant planets (e.g. Plates 2.8 and 2.51). Rings could also result from the break-up of a satellite that strayed so close to its planet that it was pulled apart by tidal forces.

Other small satellites, notably those in retrograde orbits (i.e. travelling in the opposite direction to their planet's rotation), are more likely to be captured bodies, in which case they probably began life either as asteroids or comet nuclei. The two satellites of Mars, Phobos (Plate 2.44) and Deimos, are almost certainly captured asteroids.

Neptune's only large satellite, Triton (Plate 2.23), has a retrograde orbit, which suggests that it was captured by Neptune after formation, instead of forming in a protosatellite disc. Triton is actually somewhat larger than the planet Pluto, which, so far as we can tell, resembles Triton rather closely in being a rock-ice mixture partly covered by condensed nitrogen. These may represent stunted planetary embryos, intermediate in size between the Neptune kernel on the one hand, and more distant bodies such as Smiley and Karla (Subsection 2.5.4) on the other.



(a)



(b)

Figure 2.12 Two of the largest irregular-shaped satellites in the Solar System. (a) Amalthea, an inner satellite of Jupiter (dimensions $270 \times 164 \times 150$ km). (b) Hyperion, an outer satellite of Saturn ($270 \times 240 \times 150$ km)

Summary of Section 2.6 and SAQs

- 1 The orbits of the major satellites of Jupiter, Saturn and Uranus suggest that these bodies grew out of a protosatellite disc about each planet, rather like the solar nebula in miniature.
- 2 Smaller satellites may be collisional fragments, or captured asteroids and comets.
- 3 Neptune's only large satellite, Triton, and the planet Pluto may represent stunted planetary embryos from the outer reaches of the Solar System.

SAQ 2.7 (Objectives 2.1 and 2.2) State four possible origins of planetary satellites, with examples.

SAQ 2.8 (Objectives 2.1 and 2.2) If a planetary system is forming at Beta Pictoris, what stage (if any) do you think it has reached in the sequence suggested in this chapter for the development of our own Solar System, and why?

Objectives for Chapter 2

After studying Chapter 2 (and any associated audio, video or TV material) you should be able to:

- 2.1 Give brief definitions of the terms, concepts and principles listed below.
- 2.2 Describe how the Solar System, and in particular the planets, are thought to have formed from the solar nebula, and cite some of the relevant evidence. Specifically, be able to distinguish the roles played by condensation, coagulation, planetesimal formation, and the growth of planetary embryos

List of scientific terms, concepts and principles used in Chapter 2

Term	Page	Term	Page	Term	Page
accretion	38	gravitational focusing	38	protosatellite disc	44
carbonaceous chondrite	31	kernel	40	protoSun	28
chondrite	37	massive solar nebula model	29	refractory	35
chondritic composition	32	mineral	33	runaway growth	38
chondrule	37	minimum solar nebula model	29	self-compression	41
coagulation	36	nebular theory	28	silicate	33
column mass	36	planetary embryo	38	solar nebula	27
condensation	34	planetesimal	38	turbulence	30
condensation sequence	35	primitive (composition)	31	viscous drag	30
galilean satellites	44	prograde (orbit or spin)	28	volatile	35

Chapter 3

Layering of planets

Prepared for the Course Team by Dave Rothery

Contents

3.1	Introduction	48
3.2	Internal structure of the Earth	48
3.2.1	Is the Earth layered?	48
3.2.2	Seismic evidence	49
3.2.3	Compositions of the Earth's layers	50
	Summary of Section 3.2	51
3.3	Planetary differentiation	51
3.3.1	Styles of accretion	51
3.3.2	How planetary layering arises	52
	Summary of Section 3.3 and SAQs	56
3.4	The origin of the Moon	56
	Summary of Section 3.4 and SAQ	58
3.5	Planetary heat processes	59
3.5.1	A hot topic	59
3.5.2	Heat sources	59
3.5.3	Internal heat transfer	62
3.5.4	Lithospheric heat transfer	65
	Summary of Section 3.5 and SAQs	66
	Objectives for Chapter 3	66

3.1 Introduction

In this chapter we look at what happens once a planet-sized ball of matter has come together. For example, will it be hot or cold, and will its composition vary with depth? We concentrate on the terrestrial planets, examining evidence for their present-day internal structures, and considering how these could have been brought about. We begin by looking at the internal structure of the planet for which we have the most data, the Earth. The story of how the Earth became layered is directly relevant to the other terrestrial planets and all but the smallest planetary satellites. Many of the same processes have probably affected the giant planets, which we shall consider in more detail in Chapter 7.

3.2 Internal structure of the Earth

3.2.1 Is the Earth layered?

Even though we live on the Earth we do not know exactly what it is made of. The samples of rock that geologists can collect come only from the outer 100 km or so, whereas the Earth has a radius of over 6000 km. The samples from deepest down are not from mines and boreholes (which penetrate at most only about 15 km below the present surface), but are available because the Earth is geologically active, its crust is continually being buckled by folding, and deep rock masses are thrust over shallower ones, bringing rock that formed as deep as 100 km towards the surface, where it can be exposed as a result of erosion by wind and water. Molten rock (lava) erupted at volcanoes sometimes carries up solid lumps from even deeper. However, virtually all our understanding of the deep internal structure of the Earth has to be based on *indirect* evidence.

To start examining this, let's consider the Earth's density. We can calculate the Earth's average density because we can measure its radius very precisely, and can determine its mass to several significant figures from the orbital period of artificial satellites (see ITQ 1.2).

ITQ 3.1 (a) The Earth can be approximated to a sphere of radius 6370 km. Its mass is 5.98×10^{24} kg. What is its bulk (i.e. overall average) density?

(b) The average density of the rocks that form the Earth's continents is $2.7 \times 10^3 \text{ kg m}^{-3}$ and the rocks forming the ocean floor have an average density of about $3.0 \times 10^3 \text{ kg m}^{-3}$. How do these values compare with the bulk density of the Earth, and what are the implications of this in terms of density variation with depth?

This simple argument proves that the interior of the Earth must be denser than rock near the surface. There are two factors that contribute to this. One is self-compression (Subsection 2.5.3) as a result of the inevitable increase in pressure with depth, which will affect the terrestrial planets just as it does the giant planets. The other is that the Earth's composition varies with depth, with naturally denser material concentrated towards the centre. A simple way to demonstrate this is by examining the abundances of elements in rocks near the surface. Certain elements, such as iron and nickel, are about an order of magnitude *less* abundant there than in either the Sun or carbonaceous chondrites, whereas others, such as potassium, are about an order of magnitude *more* abundant.

1 Assuming that the Earth and carbonaceous chondrites formed from solar nebula material of similar compositions, what does this tell us about the distribution of elements within the Earth?

■ Some elements (such as K) must be concentrated near the surface and others (such as Fe and Ni) must be concentrated deeper down.

In other words, the Earth must be compositionally zoned. That the structure is layered, rather than varying gradually with depth, is demonstrated by seismic evidence.

3.2.2 Seismic evidence

Most of our knowledge about the Earth's interior comes from the study of seismic waves (from the Greek *seismos*, meaning earthquake). These are vibrations generated by earthquakes or artificial explosions. If the Earth's composition and density were uniform, then the time taken for seismic waves to reach a detector would be proportional to the distance from the source. This is found not to be the case, and the data have enabled a detailed picture to be built up of how the speed at which seismic waves travel varies with depth.

Two types of seismic wave can travel through the interior of a body: compressional waves (called **P-waves**) and shearing waves (called **S-waves**). P-waves can travel through both solids and liquids, but S-waves can travel only through solids. The relationships between speed and depth for P- and S-waves in the Earth are plotted in Figure 3.1. Three major seismic divisions are recognized within the Earth, which can be seen most readily on this figure by looking at the variations in P-wave speed. There is a thin zone less than 100 km thick where P-waves travel slowly; this is the **crust**, which is typically about 25–90 km thick in continental areas, and 6–11 km thick in the oceans. At the base of the crust, P-wave speed increases sharply. This marks the top of the **mantle**, which extends down to nearly 3 000 km, at which depth there is a sharp drop in P-wave speed, marking the top of the **core**, which occupies the Earth's centre.

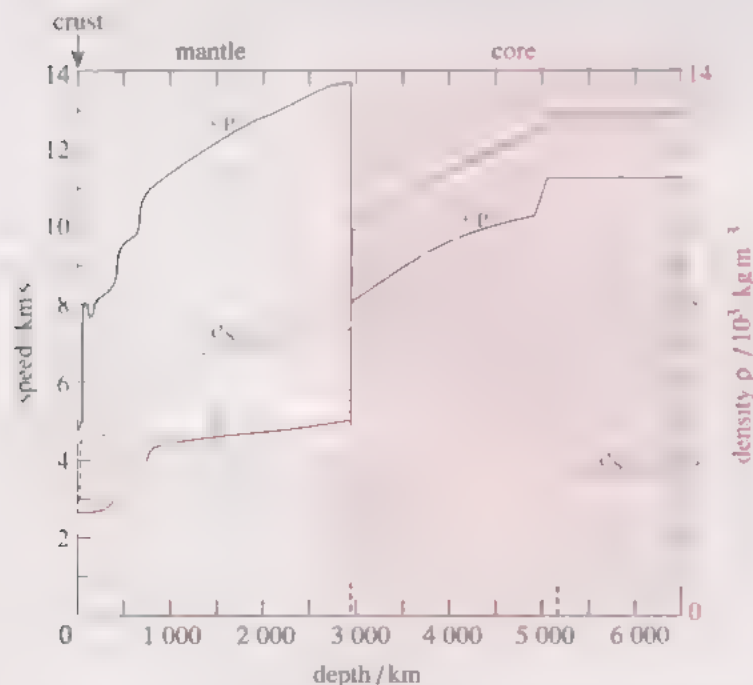


Figure 3.1 The calculated speeds of P-waves (v_p) and S-waves (v_s) versus depth in the Earth, and inferred density (ρ).

In the crust and mantle, the velocity profile of S-waves follows that of P-waves, but S-waves are not transmitted at all through the outer part of the core.

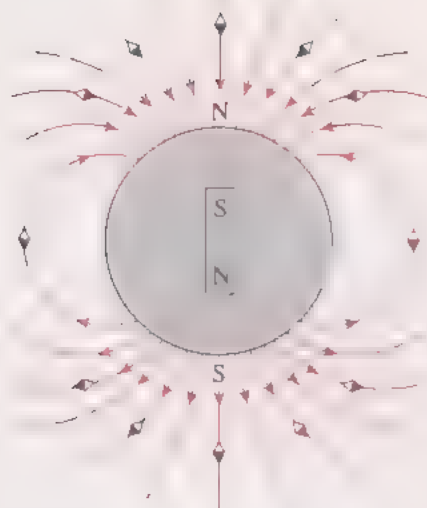


Figure 3.2 Schematic cross-section through the Earth showing magnetic field lines as deduced from the attitude of freely-suspended compass needles. The Earth has a magnetic dipole field, as if there were a giant bar magnet at its centre. The *south* pole of the Earth's internal magnet is marked at its *northern* end because this is the pole that attracts the north poles of the compass needles.

The Fe ion in Fe_2O_3 has a charge of +3 and the Fe ion in FeO a charge of +2.

Seismic velocity is a widely used term meaning the speed of seismic waves. Elsewhere you may find *seismic speed* used instead.

What does this demonstrate about the nature of the outer part of the core?

■ Remember that S-waves cannot travel through liquids, but P-waves can. The **conclusion** is that the **outer part of Earth's core must be liquid**.

Less direct means show that S-waves *are* transmitted through the *central* part of the core, revealing a contrast between this, the inner core, from the outer core that surrounds it.

ITQ 3.2 According to Figure 3.1, what is the radius of the Earth's inner core?

Further evidence that at least part of the core is liquid is provided by the Earth's magnetic field, which closely matches the field that would be produced if the Earth had a giant bar magnet at its centre; in other words, its field is that of a **magnetic dipole** (Figure 3.2). A real giant bar magnet is clearly implausible, and in fact this magnetic field cannot originate within anything solid, because at above about $1\,200^\circ\text{C}$ there is no solid that could be magnetic, and the lower mantle and core are considerably hotter than this. The only feasible way to generate such a field of the observed strength is if the core is able to conduct electricity and is in motion, perhaps as a result of convection currents.

3.2.3 Compositions of the Earth's layers

Thus the magnetic evidence is that at least part of the core is liquid, stirred up by vigorous currents, and is made of something that conducts electricity. Seismic evidence confirms that the outer part (about 95% of the core's volume) is liquid, but that the inner core is solid. The inner core is probably an alloy (mixture) of iron with about 4% nickel, similar to that in iron meteorites. The outer core is thought to be similar, but would be about 10% too dense unless mixed with a lower-density element: sulphur, potassium and oxygen have all been suggested. Iron fits the bill for the major element in the core because it is cosmically abundant (Figure 2.3), electrically conducting, and depleted in the mantle and crust.

Only the uppermost part of Earth's mantle has been sampled. Its composition is that of a rock type called **peridotite**, which has a composition equivalent to about 46% SiO_2 , 42% MgO , 7% iron oxides ($\text{Fe}_2\text{O}_3 + \text{FeO}$), and the remainder largely CaO and Al_2O_3 . Expressing composition in terms of oxides is a convenience – it does *not* mean that the minerals are simple oxides. In fact they are mostly silicates (Box 2.1), and peridotite is dominated by the minerals **olivine** and **pyroxene**, which were listed in Table 2.1.

Figure 3.1 shows that the speed of seismic waves does not increase evenly with depth within the mantle; instead there is a series of steep rises in speed, at least until 1 000 km depth is reached. These are interpreted not as changes in composition, but as depths at which the pressure forces the crystalline structure of the minerals to switch to a denser pattern.

The Earth's crust, being the outermost layer, is the least dense. It also has a different composition to that of the mantle, which explains the crust's significantly lower seismic velocities (Figure 3.1). The sharp increase in P-wave speed at the base of the crust is the most famous seismic boundary in the Earth, known as the Moho, or Mohorovičić discontinuity. There are actually two types of crust on the Earth, continental crust and oceanic crust, which, so far as we can tell, makes the Earth unique among the planets. Both types are dominated by silicates. Continental crust is typically 25–90 km thick, and its composition can be roughly expressed as 56% SiO_2 , 6% MgO , 8% iron oxides ($\text{Fe}_2\text{O}_3 + \text{FeO}$), 16% Al_2O_3 , 8% CaO , and the remainder largely Na_2O and K_2O . Oceanic crust is thinner, 6–11 km, and (as you saw in ITQ 3.1b) denser, as a result of which it is lower-lying. Its composition can be roughly expressed as 49% SiO_2 , 9% MgO ,

9% iron oxides ($\text{Fe}_2\text{O}_3 + \text{FeO}$), 16% Al_2O_3 , 11% CaO , and the remainder largely Na_2O and TiO_2 .

A summary of the Earth's layered structure is given in Figure 3.3. As you will see in the next section, it seems inevitable that most other planetary bodies have layered structures. We have no seismic evidence, except for the Moon, but variations in density with depth can be demonstrated by comparing bulk density with surface density (as in ITQ 3.1b) and from orbital studies of how their gravity fields vary with position.

Variations in gravity field with position tell us about the *moment of inertia* of a planetary body, a concept you met in Book 1, Subsection 4.4.1, from which it is possible to infer how density varies with depth.

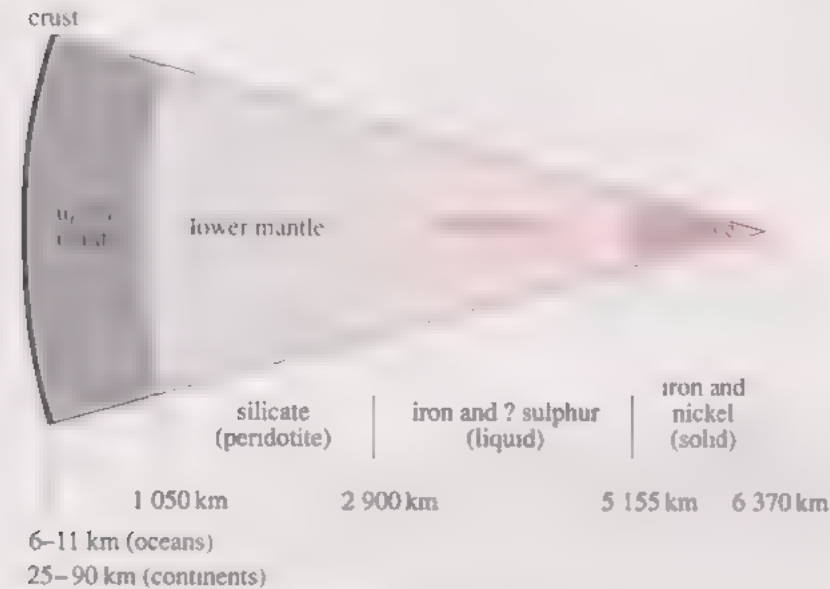


Figure 3.3 The compositional layers within the Earth. The thickness of the crust has been exaggerated for clarity.

Summary of Section 3.2

- 1 There is clear evidence from seismic data that the Earth has a layered structure.
- 2 It has an iron-rich core with a solid centre, but the outer core, which does not transmit S-waves, must be liquid. Motion in the electrically conducting outer core is responsible for the Earth's strong dipole magnetic field.
- 3 The overlying mantle consists predominantly of silicate minerals, as does the crust, though the two differ in the abundances of many elements.

There are no SAQs linked solely to this section. Instead, it is tested by parts of SAQs in Sections 3.3 and 3.5.

3.3 Planetary differentiation

3.3.1 Styles of accretion

You now know enough about the Earth's layered structure to be able consider how it arose. As you learned in Section 2.3, the first material to condense from the solar nebula at any given distance from the Sun will consist of the more refractory compounds. If a planetary body began to grow by accretion of planetesimals made from these refractory grains *before* some or all of the more volatile material had begun to condense, then the body would grow with a layered structure: the refractory compounds would be concentrated in the centre, and progressively more volatile compounds would be added concentrically around it. This style of growth is described as **heterogeneous accretion**, and is heterogeneous in the sense that the planetary body grows with a composition varying with depth.

1 Is the Earth's core more refractory than its mantle?

■ To an extent, yes. Table 2.1 shows that the iron and nickel in the core would **condense at a higher temperature than pyroxene and olivine in the mantle**. However, it is more complicated than that, because iron sulphide, a possible constituent of the outer core, is more volatile than most of the silicate minerals.

Heterogeneous accretion could happen only if the cooling of the solar nebula happened at a slow enough rate that **condensation, coagulation and even planetesimal growth** were well advanced in the case of refractory substances such as nickel-iron and pyroxene, before more volatile compounds such as alkali feldspars and then troilite (iron sulphide) condensed. It would also require that **while a planetary body was growing there was little or no mixing between planetesimals that formed simultaneously at different orbital radii**.

Can you think why?

■ Because if there was a temperature gradient in the solar nebula (Section 2.3), planetesimals that formed closer to the Sun would be more refractory than those forming further out at the same time. Mixing between these before a planetary body accreted would preclude the heterogeneous build-up of a layered structure.

An alternative model for growth of a planetary body is when material of virtually all compositions accretes onto it over a common interval. This is known as **homogeneous accretion**, because the resulting body would begin as a **homogeneous mixture throughout**.

ITQ 3.3 Does the model for the growth of terrestrial planets presented in Subsections 2.5.1–2.5.2 make heterogeneous or homogeneous accretion of the terrestrial planets the more likely?

ITQ 3.3 should have made you realize that heterogeneous accretion can probably be ruled out as a major factor in the origin of the terrestrial planets. However, the **last stage of giant planet formation, as described in Subsection 2.5.3, was heterogeneous accretion**. Their kernels may also have accreted heterogeneously (rocky material first, then the ices), depending on the balance between the rates of condensation on the one hand and coagulation, planetesimal growth and embryo growth on the other. Satellites forming within a protosatellite disc may also have grown heterogeneously.

3.3.2 How planetary layering arises

The upshot of all this is that the Earth's layered structure is probably caused by processes that occurred *after* accretion. Such segregation into compositional layers is referred to as **differentiation**. This is an inevitable consequence of the **physical and chemical properties of the constituents of a planetary body**, provided the interior is, or was once, sufficiently mobile. To appreciate how differentiation is initiated, try ITQ 3.4.

ITQ 3.4 Imagine a homogeneously accreted planet, made of chunks of nickel-iron and chunks of silicates (it does not matter how big these are). Nickel-iron is about twice as dense as silicates, and its melting temperature is some hundreds of degrees higher. What would happen to the chunks of nickel-iron if the temperature grew high enough for the silicates to melt?

The processes that led to differentiation within the Earth can hardly have failed to do likewise within the other terrestrial planets. The relative sizes of their resulting iron-rich cores and silicate (i.e. rocky) mantles are compared in Figure 3.4.

In a body composed essentially of rock and ice, like most of the satellites of the outer planets, differentiation could proceed similarly. It is doubtful if they possess enough iron for this to be a core-forming phase, or that the rock could have melted to allow the iron to separate. However, if the ice became molten or soft, the rocky fraction, being 3–4 times denser than ice, would sink towards the centre. A fully differentiated icy body therefore has a core of rock and a mantle of ice. If any icy satellite had accreted heterogeneously it would already be compositionally zoned; differentiation as described here would simply augment the separation of rock and ice. For reasons we will consider in Section 3.5, most of the major icy satellites probably were able to differentiate, ending up as shown Figure 3.5.

Thus core formation is a natural consequence of density contrasts between the predominant substances in a planet: rock versus nickel–iron in terrestrial planets and ice versus rock in icy satellites. The difference between crust and mantle is less extreme. It is essentially a result of partial melting of the mantle, which allows specific elements to be concentrated into the melt (and thus to be added to the crust) without leading to a fundamental difference in composition. As you saw in Subsection 3.2.3 the Earth's crust is richer in SiO_2 and Al_2O_3 but poorer in MgO than the mantle, though both consist essentially of silicate minerals.

As you will see in Section 3.5, there are several sources of heat within a planetary body. However, to see how most planet-sized bodies probably became hot enough for differentiation to get underway, all we need to consider is the dominant heat source during accretion, namely **accretional heating**.

What happens to the kinetic energy when one fast-moving object, such as a planetesimal, collides with another, so that they form a larger body?

■ Apart from the kinetic energy retained by any fragments that escape, it is converted to heat.

This is accretional heating. As a result of the collision, the smaller of the two bodies is likely to be completely fragmented, and the target region of the larger one will be considerably disrupted also.

It was once thought that the Earth grew by accretion of planetesimals onto a single planetary embryo. Figure 3.6 shows the consequences of accretional heating in such a situation.

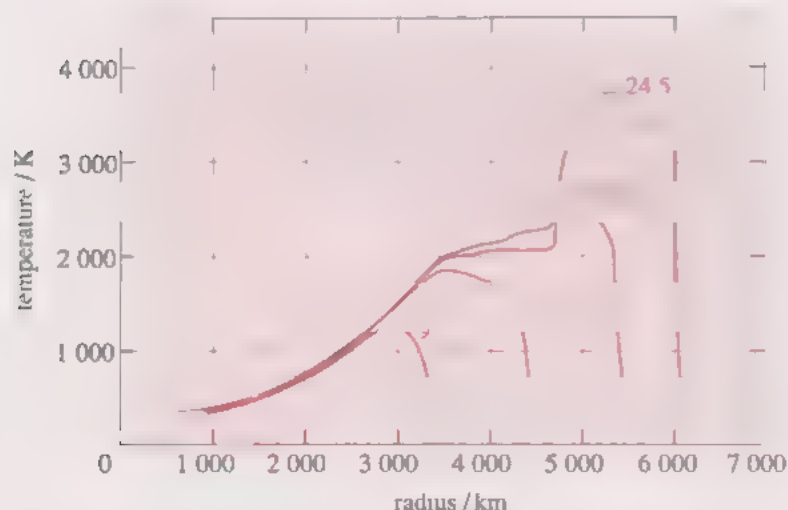


Figure 3.6 Computer simulation of the variation of temperature with depth during growth of a $1 M_{\oplus}$ planet by continual accretion of cold planetesimals (each $< 0.2\% M_{\oplus}$). Numbers refer to time in millions of years since formation. This is no longer thought to be a viable mechanism for the Earth's origin.

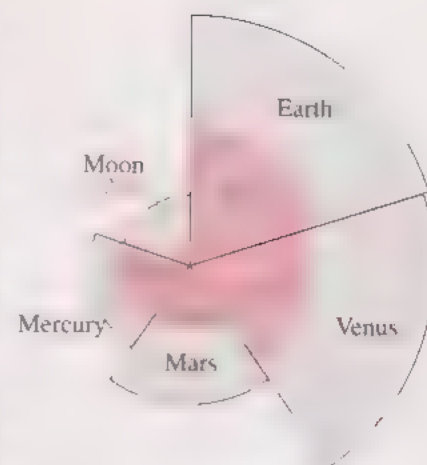


Figure 3.4 A comparison of terrestrial planet sizes and core radii, showing the percentage of the total volume occupied by the core. The models for Mercury, Venus and Mars are inferred from density and gravity-field data.

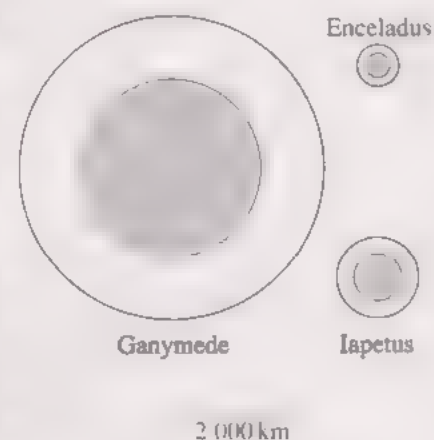


Figure 3.5 Fully differentiated models for representative icy satellites, showing their rocky cores (grey) and icy mantles. Ganymede is a satellite of Jupiter, and Iapetus and Enceladus are satellites of Saturn.

ITQ 3.5 (a) In the simulation in Figure 3.6, the outermost skin is always cold, because of rapid cooling by radiation, but why do you think the region of the planet beneath the skin is hotter than the deep interior?

(b) Can you suggest why the temperature of the region beneath the skin is higher when the planet has grown larger? (*Hint*: consider how the energy involved in collisions varies according to the mass of the target planet.)

Figure 3.6 is no longer a tenable model for growth of the terrestrial planets, if we believe the runaway growth story in Subsection 2.5.2, which says that planetesimal collisions near 1 AU would have produced planetary embryos of the order $0.1 M_{\oplus}$ (in about 5×10^4 years), and that the Earth formed by collisions between these embryos during the following 10^6 years. However, the general form of the earliest temperature–depth relationship shown, i.e. when the body had reached a radius of 3400 km (about $0.1 M_{\oplus}$), could represent that of an original planetary embryo.

In what important respect is this curve inadequate to represent a planetary embryo as described in Subsection 2.5.2?

■ According to the simulation in Figure 3.6, the 3400 km radius body took 9×10^6 years to form, which is about 200 times longer than estimated for the planetary embryos.

A planetary embryo of this size that formed more quickly than in Figure 3.6 would have a higher peak temperature just below its surface, because the same amount of kinetic energy would have been converted to heat, but with less time for heat to have been radiated away. Also, it is now thought unlikely that the planetesimals themselves would be cold. The reason (as you will learn in Subsection 8.4.2) is that studies of meteorites suggest that in the Solar System's youth a significant proportion of the aluminium present was the short-lived unstable isotope ^{26}Al (aluminium 26), created within a supernova. Radioactive decay of ^{26}Al , and unstable isotopes of other elements, would have liberated considerable quantities of heat, so that the interiors of planetesimals could become hot. This **radiogenic heating** would continue within a planetary embryo, with the result that its interior temperature would exceed that shown for the 3400 km radius body in Figure 3.6.

Radiogenic heating may have allowed the planetary embryos that went to form the Earth and other terrestrial planets to become differentiated, like small versions of the Earth today. Whether they were or not is largely irrelevant for the early history of the Earth, because of the severe disruption and accretional heating that would accompany a collision between two planetary embryos. As the smaller of the two colliding embryos would be, at most, only a few times less massive than the larger one, such a collision has become known as a **giant impact**. Figure 3.7 is a computer simulation of such an event.

What happens in a giant impact depends on the relative velocities and masses of the two bodies, but this example is fairly typical. The impact shown is between two embryos whose relative velocity is due simply to mutual gravitational attraction. In other words, the kinetic energy of the impact is derived only from the loss of gravitational energy as the two bodies came together. If the larger of the two were about $0.5 M_{\oplus}$, then the relative velocity upon impact would have a magnitude of about 9 km s^{-1} , and the kinetic energy of the impact would be about ten times that needed to melt both bodies. Thus, regardless of whether the target body was already differentiated, it would be free to do so after the impact.

In this book we often omit the atomic number of isotopes, showing only their mass number. Thus ^{26}Al is the same as $^{26}_{13}\text{Al}$.

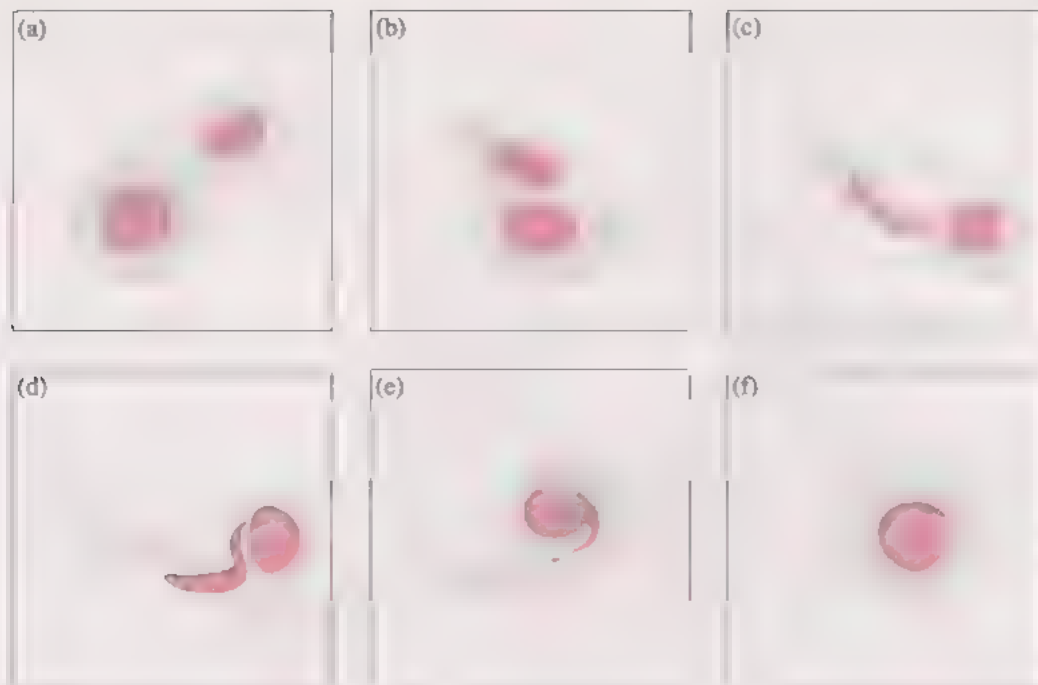


Figure 3.7 ‘Snapshot’ cross-sections showing six stages in a giant impact collision between two planetary embryos. Both are differentiated bodies, and the larger one has four times the mass of the smaller. This is a computer simulation, tracking the motion of discrete elements of each body. Iron-rich material from the cores is shown in red, and silicate-rich material is shown in grey.

The events in Figure 3.7 may be summarized as follows. In stages (b)–(c) the shock of the collision has vaporized much of the mantle of the impacting body, which is jetted, along with some of the target body’s mantle, outwards to a distance of about 2 Earth radii, from where most of it begins to fall back towards the target body as a hail of solidified debris – stages (c)–(f). Meantime, stages (d)–(e), the core of the impacting body remains more-or-less intact and ploughs its way through the mantle of the target body (which, if not already molten, would have been melted by the energy of the impact) to accrete onto the outside of its core. At stage (f) we can picture the target body as a differentiated planet with an at least partly molten iron-rich core, overlain by a molten mantle, which is termed a **magma ocean**. The surface of this magma ocean would chill and solidify because of radiation of heat to space, but this skin would continually be ruptured by vigorous convection in the underlying melt, escape of gases, and impacts by late-returning debris ejected in stages (b)–(c) and other small bodies (Plate 2.2). Escaping gases would add to any atmosphere captured from the solar nebula, or replace it with a new one if the former had been stripped away by the T Tauri wind or dispersed as a result of the shock of giant impacts. We will consider the process of outgassing as a source of planetary atmospheres in more detail in Section 6.6.

Collisions of the sort described above would have recurred several times over the 10^8 years estimated to form the Earth, and comparable histories can be envisaged for Mercury, Venus and Mars (Subsection 2.5.2). Crusts may have begun to form after each giant impact, only to be destroyed by the next. The crusts that survive today, although in some cases subsequently modified extensively by volcanism and other processes, probably began as those crystallizing out of the magma ocean after the last giant impact to affect each planet.

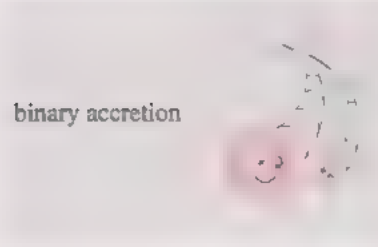
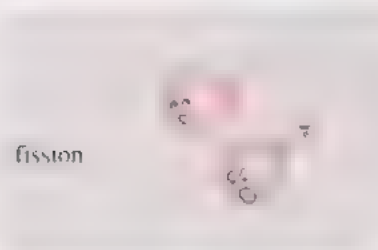
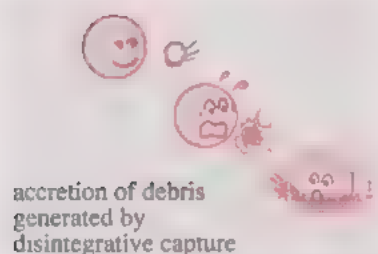
You will see in Chapter 4 that tracts of ancient crust on the Moon and several other planetary bodies record the waning phases of a period of intense bombardment. The impacting bodies were probably a mixture of locally-derived debris from giant impacts, and planetesimals and smaller fragments that had never become part of a planetary embryo.

Summary of Section 3.3 and SAQs

- 1 Radiogenic heating may have allowed planetary embryos to differentiate.
- 2 Accretional heating by giant impacts between planetary embryos will have resulted in terrestrial planets becoming differentiated early in their lives into iron-rich cores overlain by silicate mantles, regardless of the state of the original embryos.
- 3 Icy satellites that are differentiated have a silicate core, overlain by an icy mantle.

SAQ 3.1 (Objectives 3.2 and 3.3) Venus has a similar mass, size and density to the Earth, and yet its magnetic field is at least ten thousand times weaker. What can you conclude about (a) the possible existence of a core in Venus, and (b) its nature?

SAQ 3.2 (Objective 3.4) Sketch (and label) on Figure 3.6 the form of the temperature-depth curve you would expect in a 3400 km radius planetary embryo produced by collision of planetesimals over 5×10^4 years and taking account of radiogenic heating.



3.4 The origin of the Moon

The Moon's origin has long been a subject of conjecture. We look at this now, because it may well be tied in with the giant impact story. We know from data gathered by seismic detectors set up during the Apollo landings (beginning in 1969) that if the Moon has a core then it must be very small. In fact it constitutes a significantly smaller fraction by volume than the core of any other terrestrial planet (Figure 3.4).

There have been many theories put forward to explain the origin of the Moon (Figure 3.8), including that it was flung off a rapidly spinning proto-Earth (fission) or that the Earth and the Moon grew side-by-side as a double planet (binary accretion). These two theories are now untenable if we believe the model of planetesimal accretion into planetary embryos and their subsequent collision to form planets that we have followed since Section 2.4.

Another suggestion is that the Moon formed in a different part of the Solar System from the Earth, and was captured into Earth orbit later (Figure 3.8: intact capture or disintegrative capture). Apart from the difficulties of capture, analysis of lunar rock has revealed affinities with the Earth suggesting that the two bodies formed from material that had condensed in the *same* region of the solar nebula. In particular, the three stable isotopes of oxygen (^{16}O , ^{17}O and ^{18}O) occur in the Earth and Moon in relative proportions that are closely similar, whereas within meteorites (as you will see in Subsection 8.3.2) the mixture is quite different. However, the compositions of the Moon and the Earth differ markedly in other ways, as summarized in Table 3.1.

Figure 3.8 Cartoon showing some theories for the origin of the Moon.

Table 3.1 Compositional comparison of the Moon and the Earth, showing ratios between element abundances in the two bodies, averaged out over their mantle + crust. In this context volatile refers to elements such as potassium, not to gas- or ice-forming elements. Siderophile elements are those (such as nickel) that are scavenged by metallic iron, and therefore tend to have been removed from the silicates by differentiation.

	Volatile elements	Moderately volatile elements	Moderately refractory elements	Refractory elements	Siderophile elements
Moon mantle + crust	0.4–0.7	0.8	1.4	1.6–1.8	0.2–0.003
Earth mantle + crust					

Table 3.1 reveals an important trend.

□ What does it show about the differences in relative abundances of elements according to the degree of volatility?

■ Relative to the Earth, the Moon is depleted in volatile elements, but enriched in refractory elements.

This relationship favours a theory of lunar origin that gained increasing support among planetary scientists during the late 1980s, and which seems to be becoming generally accepted, though it is only fair to point out that difficulties remain in explaining some of the details. This theory is that the Moon was generated by a giant impact of the sort we discussed in Subsection 3.3.2. A computer simulation of what might have happened is shown in Figure 3.9.

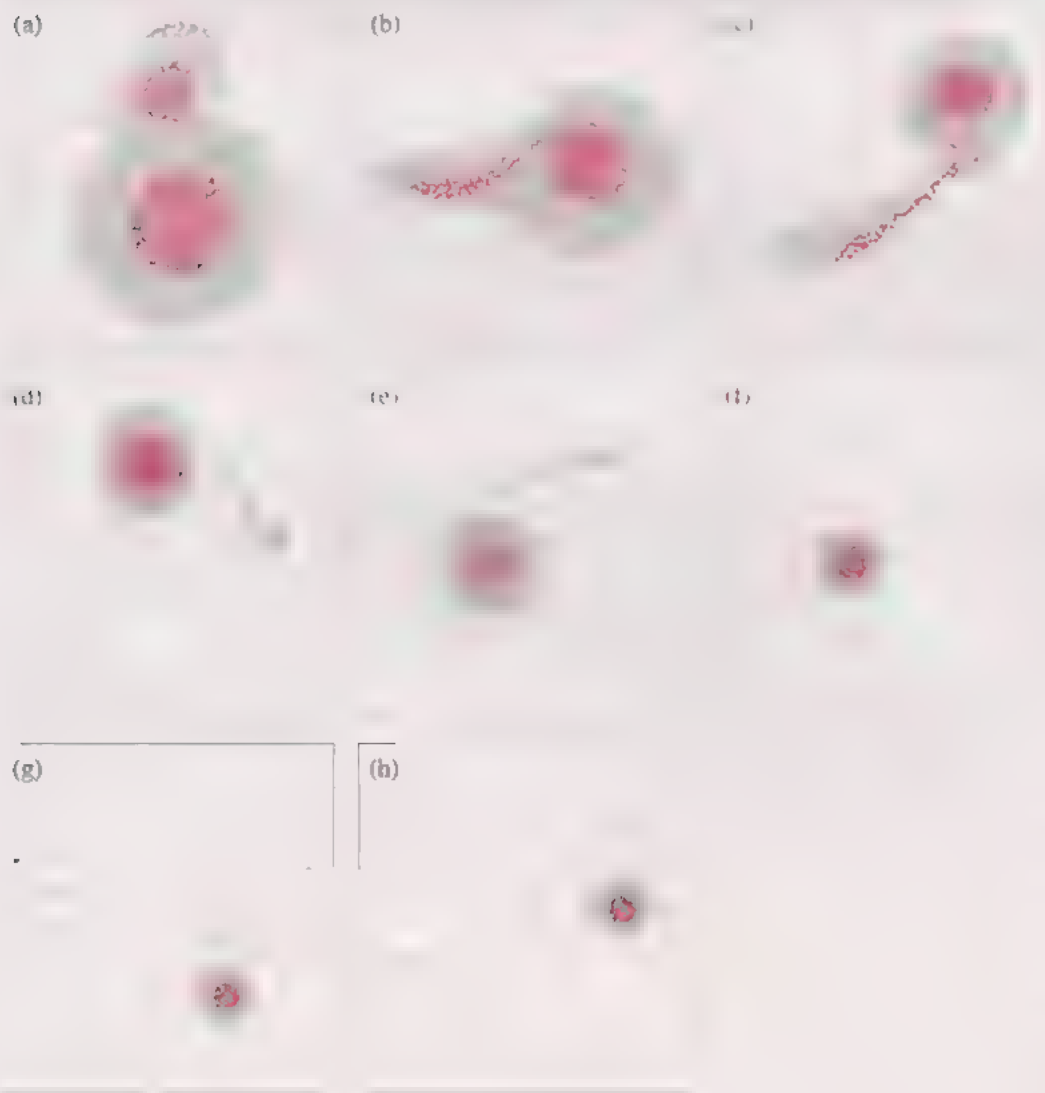


Figure 3.9 ‘Snapshot’ cross-sections in the style of Figure 3.7 showing eight stages in the impact of a Moon-forming planetary embryo onto a target body which is the proto-Earth. Iron-rich material is plotted as red dots and silicate-rich material as grey dots. In order to show more distant events, the scale of the figure is progressively reduced from (a)–(h).

The sequence shown in Figure 3.9 can be summarized as follows. In (a) a planetary embryo strikes the proto-Earth a near-grazing blow. Its core accretes onto the Earth's (stages (b)–(c)) as in Figure 3.7. Its mantle, mixed with part of the Earth's, is fragmented and vaporized, and scattered in an arc (stages (c)–(d)). Some of this material showers down on the Earth, but the rest comes together gravitationally as a clump in stable orbit around the Earth (stages (e)–(h)), which would, in time, evolve into the Moon. This was probably the last giant impact experienced by the Earth, because such a satellite would almost certainly have been destroyed by any subsequent giant impact.

This is an attractive model because it can account for many of the Moon's characteristics, not the least of which is how the Moon came to be orbiting the Earth. It also explains the Moon's anomalously small core, because the original core of the impacting body would have been incorporated within the Earth, so that the Moon's core must have formed by differentiation out of essentially mantle material.

Can you see how this model can also explain the Moon's depletion in volatile elements, relative to the Earth?

■ If the Moon grew from partly vaporized debris that recondensed after the impact, then volatile elements could have escaped.

The low abundances of siderophile elements within the Moon's mantle and crust can be attributed to their already low abundance in the essentially mantle material that formed the Moon. These abundances were reduced still further when the Moon differentiated, causing these elements to be scavenged into the newly-forming core.

The Earth–Moon similarity in oxygen isotope composition is compatible with the giant impact hypothesis too, because the colliding planetary embryos would have originated in the same region of the solar nebula. Moreover, mixing of mantles between the proto-Earth and proto-Moon during the impact would have blurred any differences.

The recognition of the role of giant impacts in planetary formation and lunar origin has led to the challenging of some of the longest-held ideas in planetary science. For example, the relatively large size of Mercury's core (Figure 3.4) was formerly used as evidence that condensation of silicates was inhibited in this part of the solar nebula because it was too hot there. However, now it seems that the large core could instead be a result of one or more giant impacts between differentiated bodies in which mantle material was dispersed into space, making the planet correspondingly iron-rich.

Summary of Section 3.4 and SAQs

The most likely explanation of the Moon's origin is that it was formed in the aftermath of a giant impact, by debris that clustered together in orbit, rather than showering down onto the Earth.

SAQ 3.3 (Objective 3.6) What are the arguments in favour of the Moon's origin in a giant impact event?

3.5 Planetary heat processes

3.5.1 A hot topic

The completion of layering within a planetary body does not mean the end of activity in the interior, in fact far from it. Virtually all the evidence of activity on the surface of a planetary body (excluding impacts and atmospheric interactions), whether it continues today or is preserved from the past, results from transfer of heat from the interior towards the outside. In this section we shall be concerned with general aspects of how heat is generated and transferred outwards. In Chapter 5 you will have the opportunity to study one consequence of planetary heat-flow, volcanism, in more detail.

Temperature profiles shaped like those in Figure 3.6 are appropriate only for very young bodies. Once the centre has warmed up, the only direction in which heat can escape is outwards, so in a mature planetary body the temperature decreases all the way from the centre to the surface.

3.5.2 Heat sources

All planetary bodies are hotter on the inside than the outside, and (for the reasons outlined in *Preparatory science*, Subsection 2.4.2) this heat is transferred outwards. At the surface, heat is lost by radiation. The mean surface temperature is a consequence of a balance between the rate of this heat loss, on the one hand, and the rates of heat flow from the interior and of heating by absorption of solar radiation, on the other. If there is an atmosphere it can trap heat, so the surface temperature can be higher than this theoretical equilibrium (as you will see in Subsections 6.4.1 and 6.4.2), but the basic logic remains the same.

Heat is escaping from the Earth's interior to its surface at a global average rate of 0.08 W m^{-2} , and (based on far fewer measurements) the global average lunar heat flow is estimated at about 0.015 W m^{-2} . In this subsection you will see how the internal heat of these and other solid planetary bodies is generated.

Accretional heat

You have already seen that the conversion of kinetic energy to heat during giant impacts is more than enough to melt bodies comparable in size to the terrestrial planets. The collision illustrated in Figure 3.7 would liberate about $4 \times 10^7 \text{ J}$ per kilogram of impactor (about $2.4 \times 10^{31} \text{ J}$ for $0.1 M_{\oplus}$). This energy would end up as heat, except for the kinetic and gravitational energy of any debris that escaped ultimate accretion onto the planet (probably a tiny fraction). Even bombardment of a young planet by swarms of planetesimal-sized objects could provide enough energy to initiate a magma ocean. All planetary bodies must have experienced accretional heating while they were being formed, and heat will have been added near the surface by the smaller impacts that occurred later. Today this is a negligible heat source, because most available debris was swept up by the planets long ago.

Core formation

Another source of heat that may have played a significant role in the early history of a planetary body is the core-forming process during differentiation.

If an undifferentiated planetary body differentiates, with denser material segregating inwards to form a core, why is energy released?

■ There is a loss of gravitational energy, which will be converted to thermal kinetic energy, in other words to heat (Book 1, Subsection 3.2.3).

The total energy available from **core formation** by this means within the Earth would be some $1.3 \times 10^{30} \text{ J}$. According to the model for the Earth's formation that we have described, much of this heat would have been generated spasmodically, each time the core was enlarged by a giant impact. However, it is

possible that core growth is continuing (at a far slower rate) within the Earth today, and similarly within other planetary bodies.

Apart from in the giant planets where core formation may still be a significant heat source (Subsection 7.2.1), accretional heating and heating by core formation must be considered as essentially primordial processes. If these were the only heat sources to have operated within the smaller planetary bodies, there would have been time for most of this heat to have leaked away. To understand why some planetary bodies remain geologically active, we must consider processes of heat generation that continue today.

Radiogenic heating

You have already encountered radiogenic heating (Subsection 3.3.2), in the form of energy released by radioactive decay of ^{26}Al , a short-lived isotope that has long since vanished. This and other short-lived isotopes probably liberated enough heat to differentiate planetary embryos even before their first giant impact collision. The same heat source may explain the heating of meteorite parent bodies, of which more in Section 8.4. Within planetary bodies today, practically all radiogenic heating comes from the decay of long-lived isotopes of just three elements: uranium (U), thorium (Th) and potassium (K). Most minerals in rock contain small amounts of these isotopes, and their relative importances in the Earth are summarized in Table 3.2. Decay of their nuclei to more stable forms releases heat, as outlined in Box 3.1.

Table 3.2 Isotopes important as radiogenic heat sources today

Isotope	Half-life/ 10^9 yr	Present rate of heat generation in the Earth $10^{-9} \text{ mW kg}^{-1} a$	Abundance today Abundance when Earth formed
^{235}U	0.71	0.04	0.013
^{238}U	4.50	0.96	0.50
^{232}Th	13.9	1.04	0.80
^{40}K	1.30	2.8	0.092

^a Values are per kg of the Earth, not of the isotope concerned.

Box 3.1 Radioactive decay

There are four radioactive isotopes whose decay produces significant amounts of heat and that (thanks to their long half-lives) are still abundant in the Earth: uranium-235 and -238 (^{235}U and ^{238}U), thorium-232 (^{232}Th) and potassium-40 (^{40}K); see Table 3.2.

As an example, ^{235}U (the parent isotope) decays to ^{207}Pb (its daughter isotope). Energetic particles (in this case α particles) are emitted, and their collision with other nuclei causes heating. After 0.71×10^9 years, half of the original ^{235}U nuclei will have decayed to ^{207}Pb (this is the definition of the term half-life, see also Book 1, Section 4.3.3), and the remainder continues to halve every 0.71×10^9 years.

In a mineral grain that originally contained ^{235}U but no ^{207}Pb the age of the grain can be calculated from the proportions of these two isotopes present today, as you can see graphically in Figure 3.10. Comparing ratios of parent and daughter isotopes in this way is the basis of **radiometric dating**.

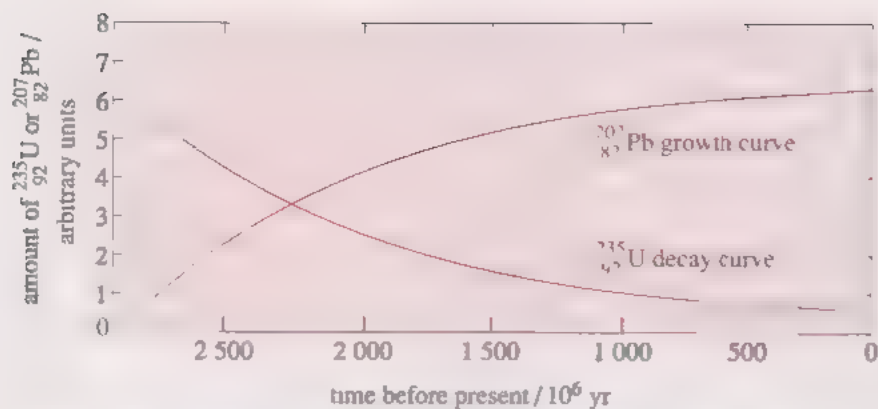


Figure 3.10 The changing proportions of ^{235}U and ^{207}Pb over time. The decay curve for the parent isotope, ^{235}U , is a mirror image of the growth curve for the daughter isotope ^{207}Pb .

With the possible exception of potassium, the elements with important long-lived radioactive isotopes are concentrated into the silicate part of the Earth, and especially in continental crust, where the rate of radiogenic heat production per kg is about a hundred times that in the mantle. The mass of the Earth's mantle is about a hundred times that of the crust, so radiogenic heating is split about equally between crust and mantle.

Decay of long-lived isotopes provides sufficient heat to keep the Earth and Venus geologically active. The smaller terrestrial planets were active in the past when these isotopes (and some now-vanished short-lived ones) were more abundant, but by now they are too cool. This is because the smaller the body, the greater its surface area in proportion to its volume, and so the faster it can radiate away the heat generated per unit volume.

Tidal heating

Another source of heat that may be significant within a planetary body is **tidal heating**, which is a result of tidal distortion of the shape of the body. Just as Earth's oceans experience tides caused by the gravitational attraction of the Moon and the Sun, so the solid Earth itself is distorted by these forces, producing tides with a maximum amplitude at the surface of about 0.1 m.

Tidal heating is the main subject of video sequence 5, which you should now watch.

As shown in the video sequence, tidal heating is particularly important among the satellites of the outer planets, the most spectacular example of which is Io (Plates 2.26-2.28).

Having watched video sequence 5, you should appreciate the importance of orbital resonance within satellite systems. This prevents satellites' orbits from becoming circular so that strong tidal forces continue to affect their tidal bulges even though the satellites are in synchronous rotation. Tidal heating within the Earth is different, the Earth rotates much more rapidly than the Moon orbits it, so tidal bulges have to race around the globe almost once per day to keep pace with the Moon. Exactly where within a planetary body most of the tidal heat is generated depends on its internal properties, which even for the Earth are poorly known at high pressure; however, it is a reasonable assumption that the effect is greatest in the crust and upper mantle.

ITQ 3.6 The present rate of heat generation within the Earth by tidal distortion is about $3 \times 10^{19} \text{ J yr}^{-1}$. Given that the Earth's mass is approximately $6 \times 10^{24} \text{ kg}$, what is the rate of tidal heating per kg, expressed as mW kg^{-1} , and how does this compare with the rate of radiogenic heating? (Remember $1 \text{ W} = 1 \text{ J s}^{-1}$.)

ITQ 3.6 shows that tidal heating within the Earth is between one and two orders of magnitude less than the rate of radiogenic heating (Table 3.2), so it probably has little influence on the Earth today. In the past, the rate of tidal heating may have been greater, but then so was the rate of radiogenic heating.

One result of the tidal interaction between the Moon and the Earth has been to slow down the Earth's rotation. By counting daily growth lines in certain kinds of fossil organisms, it appears that there were about 400 days per year 400 million years ago (implying a day-length of around 21.5 hours).

□ If the Earth's rotation has slowed down, what has happened to its angular momentum?

■ Some of it must have been lost.

However, angular momentum must have been retained within the Earth-Moon system, and in fact what the Earth has lost has been gained by the Moon in terms of orbital motion. As a result of this effect (which is still continuing), the Moon is receding from the Earth at a rate of about 3 cm per year.

3.5.3 Internal heat transfer

We must now consider the effects of heat within a planetary body. At the start of the previous subsection it was stated that heat generated in the interior is transferred to the surface, whence it is radiated away to space (with or without the intermediary of an atmosphere). But how does heat get from the interior to the surface?

□ What mechanisms of heat transfer have you met within stars?

■ In Book 1, Subsection 1.5.5, we described convection and radiation as important mechanisms, and stated that conduction plays a very minor role.

We can rule out radiation as a significant mechanism in a planet, because their interiors are opaque and the temperature is too low. Here heat transfer is by conduction and convection. In solids the most familiar mechanism of heat transfer is **conduction**. You can experience conduction for yourself by holding a poker in a fire, and feeling how the end in your hand gradually gets too hot for comfort! What happens is that heat transfer takes place on an atomic scale, by transfer of internal energy as described in *Preparatory science*, Subsection 2.4.2.

What about convection within planets?

What property must a material possess to enable heat to be transferred by convection?

■ It must be fluid. [*Comment.* The term 'fluid' describes any substance capable of flow, and is not restricted just to liquids and gases.]

In Subsection 3.2.2 you saw that the Earth's outer core must be liquid, because it does not transmit seismic S-waves, and that convection currents may be occurring there. As the Earth's mantle is solid, you might think that this leaves conduction as the only viable means for outward transfer of heat from the core, and of heat generated within the mantle. However, this is not the case. Silicate rock will flow convectively *as a solid* if the temperature and pressure are high enough and the thermal gradient (the rate at which temperature increases with depth) is steep enough. This phenomenon is known as **solid-state convection**. It is a slow process, with rates of flow no more than a few centimetres per year, but nevertheless it is more efficient than conduction as a means of heat transfer except through the outermost shell of a solid planetary body (which is too cold and rigid to convect at all).

Various lines of evidence show that much of the Earth's mantle must be mobile. One of the first to be appreciated is that Scandinavia is rising at a rate of

You may find flow hard to visualize in a solid. Pitch is clearly a solid: it is capable of transmitting both P- and S-waves, and if you hit it sharply with a hammer it will fracture. However, a block of pitch standing at room temperature will spread out, or flow, under its own weight.

several millimetres per year, rebounding in response to the disappearance of a thick ice sheet that covered it during the last glaciation, 15 000 years ago. This rise (and the original depression when the ice was loaded onto the region) could occur only if material were able to flow at depth. The weak, deformable part of the Earth became known as the **asthenosphere** (from the Greek *astheno*, meaning weak) as distinct from the overlying rigid shell known as the **lithosphere** (from the Greek *lithos*, meaning rock) which is about 100 km thick.

The importance of the distinction between the lithosphere and the asthenosphere (which is *not* the same as the difference between the crust and the mantle) became clear in the 1960s with the development and acceptance of the theory of **plate tectonics**, which recognizes that the Earth's lithosphere is broken into rigid plates that are able to slide around at about 10 cm yr^{-1} because of a weak layer beneath. Figure 3.11 summarizes the elements of plate tectonics, with which you may already be familiar if you have studied the *Science Foundation Course*.

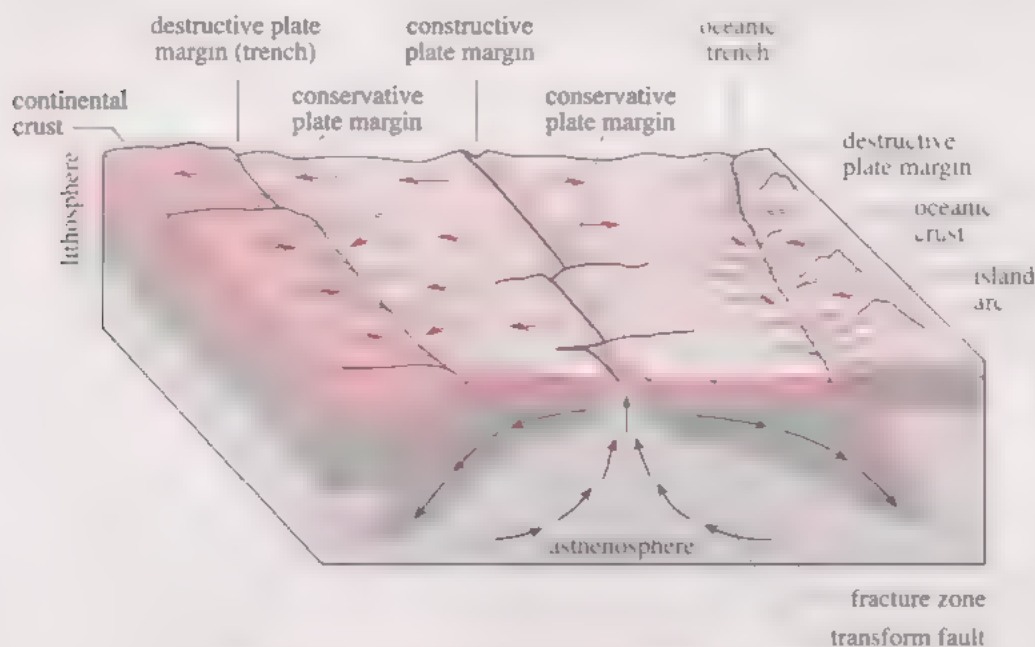


Figure 3.11 Simplified block model (vertical scale exaggerated) showing the elements of plate tectonics in the Earth. The lithosphere is rigid, and comprises the crust and the uppermost part of the mantle, whereas the asthenosphere is weak and convecting, and comprises the underlying mantle. Lithospheric plates move apart at constructive plate margins (where they are continually added to by volcanism), slide past each other at conservative plate margins, and are destroyed at destructive plate margins (where one plate is forced down below another, and is assimilated into the asthenospheric mantle at depth).

In the Earth, plate motion is facilitated by the particularly weak nature of the upper zone of the asthenosphere.

Look back at the seismic speeds plotted in Figure 3.1. Can you see anything noteworthy anywhere between 0 and 500 km depth?

■ There is a dip in both v_p and v_s at about 150 km depth, whereas throughout the rest of the crust and mantle speeds increase with depth.

This 'low-speed layer' coincides more or less with the upper part of the asthenosphere and is thought to be a result of partial melting: that is to say, there is a few per cent of melt in the interstices between crystals. Clearly if motion between crystals is lubricated by melt then convection will be easier. However convection probably continues throughout the full thickness of the mantle. The rate of convection within the deep mantle, where the viscosity is probably several times greater than in the low-speed layer, is probably about $0.1\text{--}1.0 \text{ cm yr}^{-1}$,

According to some definitions, the term asthenosphere is restricted to the outer zone of the convecting mantle. The more viscous, more slowly convecting, deep mantle is then distinguished by the term 'mesosphere'. In this Course we use asthenosphere to embrace the whole of the mantle beneath the lithosphere.



Figure 3.12 Convection within the Earth's mantle

whereas at the top of the asthenosphere it must match the speed of plate motion, at around 10 cm yr^{-1} . The general nature of convection patterns in the Earth's mantle is indicated in Figure 3.12.

The lithosphere, being rigid, is unable to convect, so out of the three methods of heat transfer noted above, heat can pass through it only by conduction. This provides the basis of general definitions, which we can extend to solid planetary bodies where we have no seismic data:

The lithosphere is an outer, rigid layer within which heat can be transferred by conduction but not convection, whereas the asthenosphere is an underlying weaker layer that is capable of flow and within which the principal means of heat transfer is convection.

Ice, just like silicate rock, will flow by solid state convection under the right conditions (which might occur in the *interior* of an icy satellite), but it is rigid at the low *surface* temperatures prevailing in the outer Solar System. Therefore these definitions can be applied just as well to icy satellites as to terrestrial planets.

The thickness of a planetary body's lithosphere must increase as the internal temperature declines, and the top of the convecting asthenosphere retreats inwards. A body within which too little heat is generated will be lithospheric throughout. The probable thicknesses of lithospheres in several planetary bodies are contrasted in Figure 3.13. Only in those bodies whose lithospheres are relatively thin is molten material likely to be able to reach the surface by means of volcanism, and similarly a lithosphere must be thin if it is to be cracked and buckled as a result of convection within the asthenosphere.

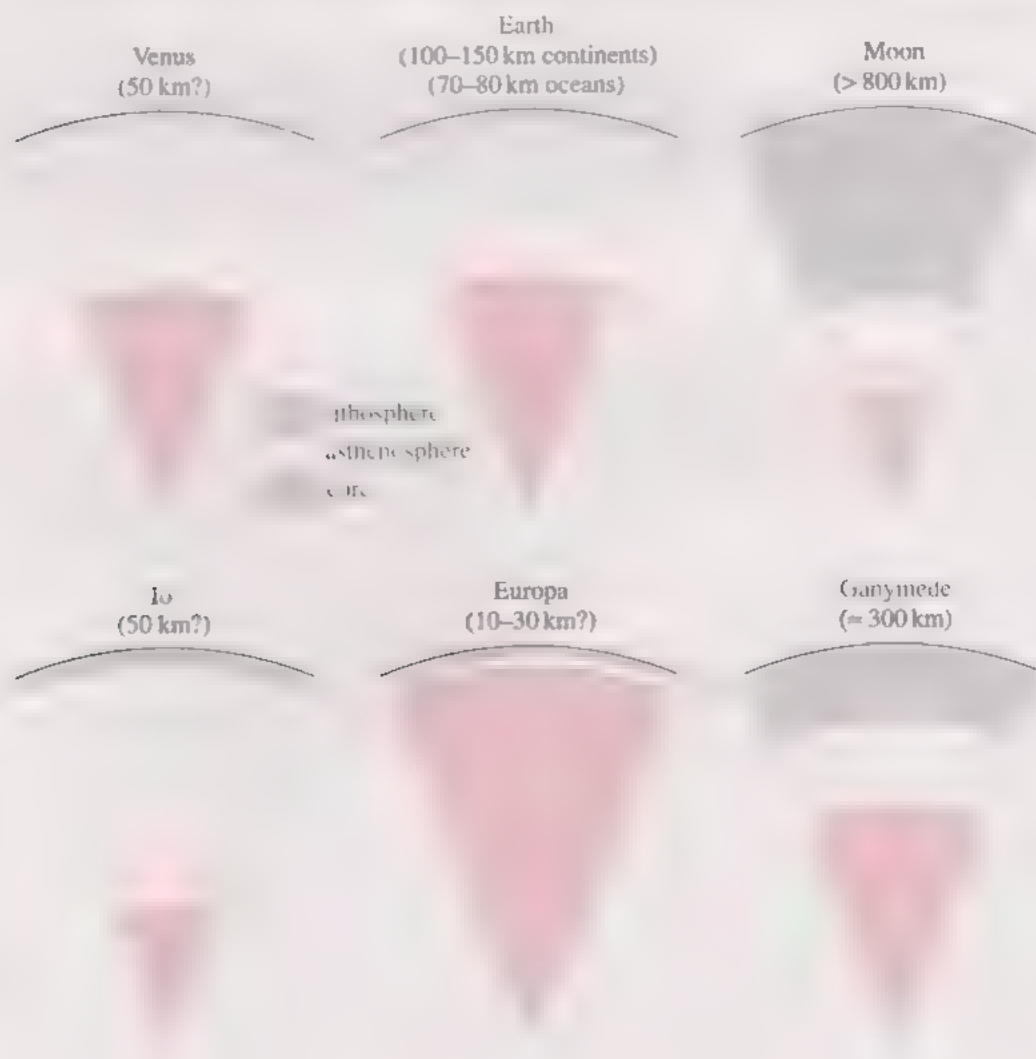


Figure 3.13 Lithospheric thicknesses in a range of planetary bodies, drawn as if they were the same size.

3.5.4 Lithospheric heat transfer

In the Earth and comparable planetary bodies, heat is carried outwards to the base of the lithosphere mostly by convection in the asthenosphere. This heat must pass through the lithosphere in order to reach the surface, along with the heat generated within the crust and the lithospheric part of the mantle. By definition, the lithosphere is a conductor, rather than a convector of heat. In the Moon today, which has a very thick lithosphere, all the heat arriving at its surface does so by means of conduction. However, there are ways other than conduction of carrying heat through a lithosphere that operate on geologically active bodies such as the Earth. These are well worth considering, because they allow useful comparisons to be made between planetary bodies.

One way of transferring heat is if hot, molten material (magma) can pierce the lithosphere and rise through it. Whether it spreads out over the surface as a lava, or cools and crystallizes at some depth below the surface as an intrusion, the effect is the same: heat is carried from a deeper level to a shallower one. Transport of heat in this way is known as **advection**. In any planetary body with active volcanism, some heat transfer must occur advectively.

There is a third process of heat transfer involving the lithosphere that it is useful to distinguish on the Earth. This is **plate recycling**, of which you can see the basis in Figure 3.11. Old, cold parts of lithospheric plates are dragged down (subducted) into the deeper mantle at destructive plate margins, thereby adding cold material to the asthenosphere, whereas volcanism along constructive plate margins adds hot material (derived from the asthenosphere) to the newly forming edges of the plates. Once again, there is a net outward transfer of heat.

The relative importances of these three methods of lithospheric heat transfer in different planetary bodies are indicated on the triangular plot shown in Figure 3.14. Plates 2.3–2.28 show a variety of planetary bodies (beginning with ‘dead’ ones and running through to those active today) on which the degree of volcanism and deformation of the lithosphere appears to be a consequence of different lithospheric thicknesses. In Chapter 5 we will examine in more detail why volcanism occurs on some planetary bodies, and see how the different styles of volcanism can be related to both their overall heat-flow and their composition. First though, in Chapter 4, we will look at cratering.

In plate recycling, heat is carried by wholesale creation and destruction of the lithosphere, whereas in advection heat is carried by hot material rising into or through the lithosphere.

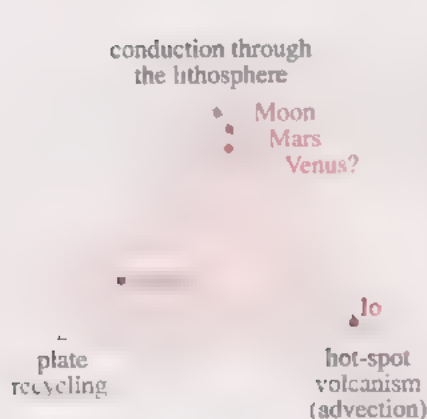


Figure 3.14 Relative importances of conduction, advection and plate recycling in the lithospheres of several planetary bodies. The closer a point lies to an apex, the greater the relative contribution of that process to heat transport through the lithosphere. Apart from the Earth, the points shown represent best guesses. These values are for the present day. In the past, when smaller bodies had thinner lithospheres, the relative importance of conduction would have been less. On the Earth, there is a lot of volcanism associated with the creation and destruction of lithosphere at plate boundaries. For the purposes of this plot, heat carried by those volcanoes is regarded as part of the plate recycling effect; only the heat carried outwards by volcanoes at hot-spots distant from plate boundaries is counted as advection.

Summary of Section 3.5 and SAQs

- 1 Heat is generated within a planetary body today mainly by radioactive decay of ^{235}U , ^{238}U , ^{232}Th and ^{40}K within silicates (radiogenic heating), and by tidal interactions when two or more bodies interact gravitationally. Accretional heating, decay of short-lived isotopes (such as ^{26}Al), and core formation were important in the past.
- 2 An asthenosphere is a zone where temperature, pressure and temperature gradient are high enough to allow material to flow, even in the solid state, so that heat is transferred outwards mainly by convection.
- 3 A lithosphere is rigid and cannot convect; instead heat is carried through it by conduction. On geologically active worlds heat can also be transported through the lithosphere by advection, or be brought to the surface by wholesale recycling of the lithosphere. The Earth's lithosphere includes the crust and the uppermost part of the mantle.
- 4 These processes of heat transfer can occur equally well in ice as in silicate rock.

SAQ 3.4 (Objective 3.4) Three of the major satellites of Saturn are Mimas (pronounced 'miv-mass'), Enceladus ('en-sel-adus') and Dione ('die-oh-nee'). Their orbital periods are 0.942, 1.370 and 2.737 days respectively. Does any pair of these satellites show orbital resonance, and if so, how would this contribute to heating? [*Hint*, work out the ratios of the orbital periods to two significant figures.]

SAQ 3.5 (Objectives 3.4 and 3.5) (a) What changes will there be over time in the thickness of the lithosphere of a planetary body within which the dominant heat source is radiogenic heating?

(b) Does the same apply to bodies where tidal heating is dominant? Give reasons.

SAQ 3.6 (Objectives 3.2 and 3.5) Which of the following layers in the Earth are defined seismically, and which by other means? What do the differences between these layers represent? How are they defined if not seismically? Asthenosphere, core, crust, lithosphere, mantle.

Objectives for Chapter 3

After studying Chapter 3 (and any associated audio, video or TV material) you should be able to:

- 3.1 Give brief definitions of the terms, concepts and principles listed below.
- 3.2 Cite the evidence that the Earth, and by implication other planetary bodies, is internally layered.
- 3.3 Describe the reasons why, and the processes by which, layering of planetary bodies is likely to have developed.
- 3.4 Describe the alternative sources of heat within a planetary body.
- 3.5 Describe the alternative mechanisms by which heat is transferred from the interior of a planetary body, and in particular how it is transferred through the lithosphere.

3.6 Cite some of the evidence that the Earth–Moon system is the result of a giant impact, and describe what may have happened during such an event.

List of scientific terms, concepts and principles used in Chapter 3

Term	Page	Term	Page	Term	Page
accretionary heating	53	heterogeneous accretion	51	plate recycling	65
advection	65	homogeneous accretion	52	plate tectonics	63
asthenosphere	63	lava	48	radiogenic heating	54
conduction	62	lithosphere	63	radiometric dating	60
core	49	magma ocean	55	S-waves	49
core formation	59	magnetic dipole	50	seismic waves	49
crust	49	mantle	49	solid-state convection	62
differentiation	52	P-waves	49	tidal heating	61
giant impact	54	peridotite	50		

Chapter 4

Cratering

Prepared for the Course Team by Peter Francis

Contents

4.1	Introduction	69
4.2	Identifying impact craters	70
4.2.1	Diagnostic features of impact craters	71
4.2.2	Characteristics of pristine lunar craters	73
4.3	Shapes of craters	75
4.3.1	Effects of target materials	76
4.3.2	Effects of scale	78
	Summary of Sections 4.2 and 4.3 and SAQs	81
4.4	Craters as chronometers	83
4.4.1	Plotting crater size-frequency distribution curves	85
4.4.2	Using craters as chronometers	87
4.4.3	A calibrated cratering curve	88
4.4.4	Comparisons with Mars	89
	Summary of Section 4.4 and SAQs	91
	Objectives for Chapter 4	93

4.1 Introduction

In completing one orbit around the Sun, the Earth travels about 942 million km in 365 days. Its orbital speed is about 30 km s^{-1} . All the other objects in the Solar System are moving at similar speeds. If their orbits intersect, their closing velocities can be huge, with enormous amounts of kinetic energy being liberated in collisions. This explains why **impact cratering** is by far the most widespread process shaping the surfaces of solid bodies in the Solar System. No object, large or small, is exempt. On the Moon, impact structures of all sizes are ubiquitous, including huge basins over 1 000 km in diameter, visible to the naked eye from the Earth. On many of the smaller satellites and asteroids in the Solar System, cratering is the *only* process to have left a mark since their formation. Craters are so characteristic of extraterrestrial bodies that when cartoonists want to caricature alien life, they almost always draw bug-eyed monsters peering above crater-covered landscapes.

As we hope you discovered through your project work, if you study the Moon with binoculars when the Sun's rays are shining obliquely on it, so that shadows are well formed, you will see a battered surface, pitted with craters. If you were to view the Earth from the Moon, however, you would see scarcely any evidence of impact cratering at all. Only one crater, the magnificent ring of Lake Manicouagan in Canada (Figure 4.1) might be discernible in good conditions.

If impact cratering is indeed an ubiquitous process, why does the Earth show so little evidence of it?

■ On Earth, a range of geological processes has been at work, wiping out the traces of early massive bombardment.

Impacts have played an important, perhaps decisive, role in the evolution of life on Earth. Some scientists believe that life itself – however it originated – could not thrive on Earth until the period of heavy bombardment ended, about 4 billion years ago. While the record on the Earth of that bombardment is irretrievably lost in the abyss of time, there is *direct* evidence of a much more recent catastrophic impact.

For decades, palaeontologists were puzzled by the abrupt mass extinctions of the dinosaurs and many other animal and plant species 65 million years ago. Recent work has shown, beyond a reasonable doubt, that a major impact took place at the time of the extinction event. A buried crater at least 180 km in diameter at Chicxulub (pronounced 'chuck-sulub'), on the Yucatan peninsula of Mexico, is of exactly the right age, and is now generally accepted as the impact site.

Discovery of the Chicxulub crater does not entirely resolve the mystery of the extinctions. It is not clear that all the extinctions were instantaneous – some species appear to have been already in decline before the impact event. It also leaves unresolved the precise mechanism of how the impact could have led to the extinctions of so many species.

Whatever the case, the fact that the Earth will inevitably experience other substantial impacts in the future means that they are of more than purely academic interest. This chapter should give you an understanding of the cratering processes, and enable you to describe the impact craters you would expect to find on an unknown planetary body, given only basic information about its size, mass, density and orbital parameters. You should also understand the principles involved in crater size-frequency distribution statistics, so that you can use these to estimate the ages of planetary surfaces.

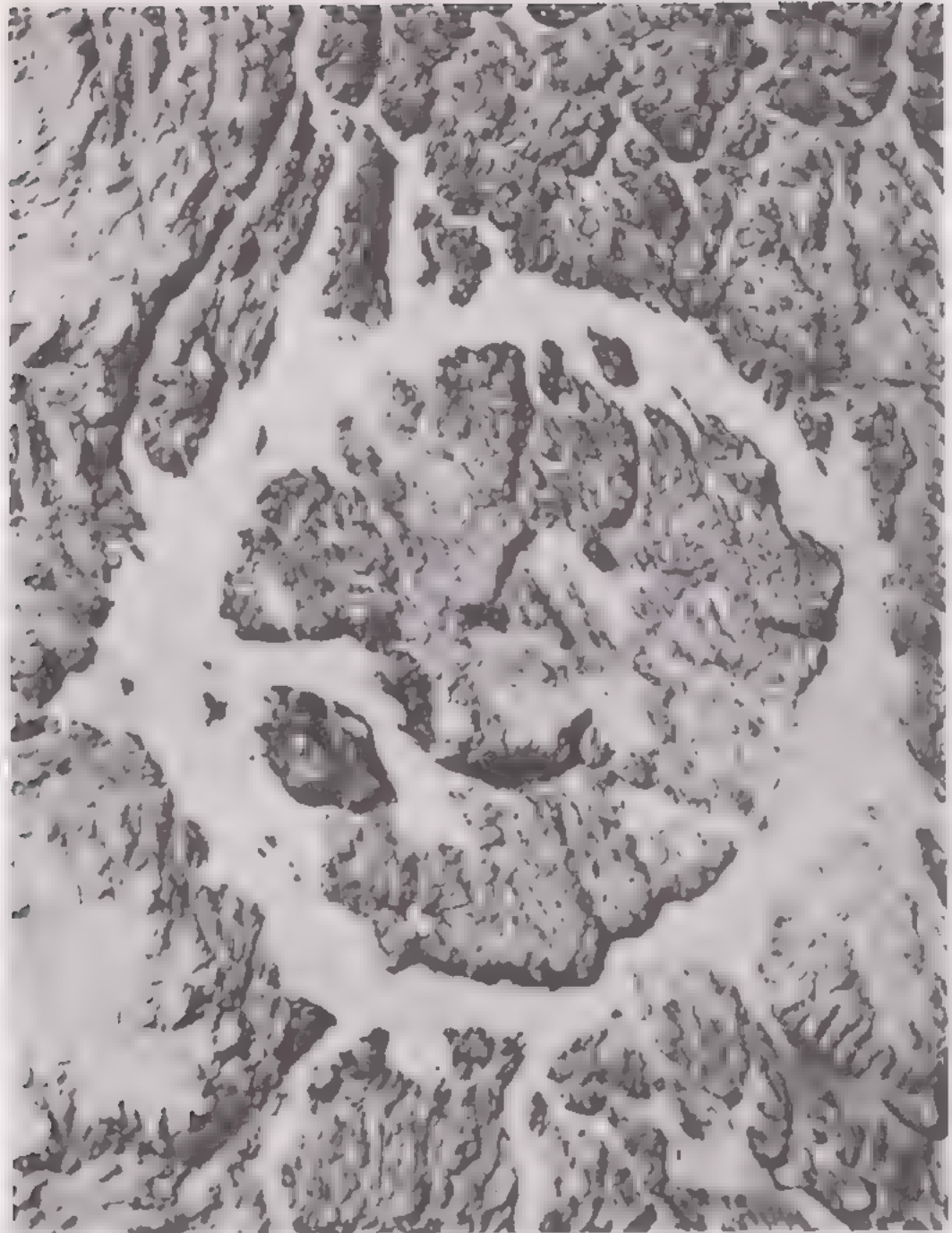


Figure 4.1 Lake Manicouagan, a 65 km diameter impact crater in Quebec, Canada, believed to have formed about 200 million years ago. The concentric structure of the crater is highlighted by the waters of the lake, formed by the Daniel Johnson dam across the Manicouagan river (bottom).

4.2 Identifying impact craters

In astronomical terms, the Moon is extremely close to the Earth, and easily observed. Thus, the topography of its craters was minutely studied by many generations of observers, who measured, sketched and photographed thousands of them. None the less, for decades before the Apollo lunar landings (1969–1972), there was no consensus about how the craters formed. One school of thought, initiated in 1893 by the American geologist G. K. Gilbert (1843–1918), believed that they were indeed of impact origin. Another school argued that they were volcanic, failing to recognize that lunar craters are manifestly quite unlike terrestrial volcanic craters in scale, shape or structure. This controversy shows

how scientific debates may remain interminably deadlocked, until some new factor tips the argument one way or another. The Apollo missions provided the impetus to resolve the crater controversy, though most scientists accepted the impact hypothesis before the first landing actually took place. Open University students may be gratified to know that some of the most persuasive arguments for the impact origin of lunar craters came not from professional scientists at all, but from a talented American business man, R. B. Baldwin (1912–), who studied the Moon in his spare time, and published an influential book entitled *The Face of the Moon* (1949).

Given that terrestrial volcanic craters are poor analogues of lunar craters, suggest one reason why the volcanic hypothesis was popular amongst terrestrial geologists.

■ Craters produced by volcanic eruptions were familiar; part of geologists' everyday experience.

Early studies of impact cratering, therefore, were clouded by the 'geocentric' education of scientists: impact events did not form part of their world view. No significant impact crater has been formed on Earth in historic times. Similar problems overshadowed the more recent debate over the mass extinctions 65 million years ago: many conservative paleontologists were reluctant to accept the possibility of catastrophic impact events, which were beyond their experience.

You should now watch video sequence 6, which relates to the following material. (You might also find it helpful to view video sequence 6 again later, after completing this chapter.)

4.2.1 Diagnostic features of impact craters

Of the 100 known impact craters on the Earth, **Meteor Crater, Arizona** (Figure 4.2) has contributed more than any other to the understanding of impact cratering. Perhaps surprisingly, it was not universally accepted as an impact crater until the 1960s. Its impact origin was finally demonstrated in the 1960s by the American geologist Eugene Shoemaker (1928–), who has made a life study of cratering.

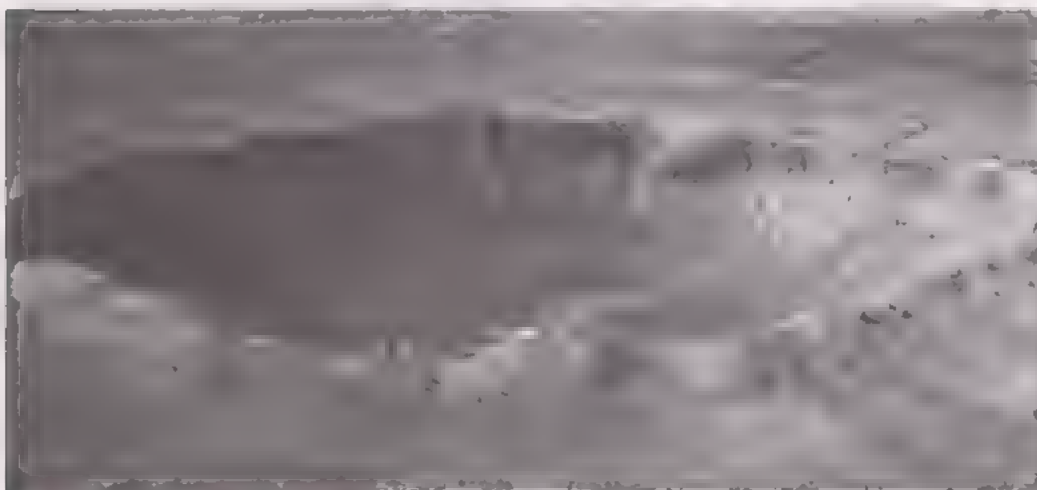


Figure 4.2 Oblique aerial view of Meteor Crater, Arizona, USA, showing its raised rim and the subdued, surrounding hummocky ejected deposits (foreground). Meteor crater is 1.2 km in diameter and 200 m deep, with a raised rim 50–60 m high, and was formed about 40 000 years ago.

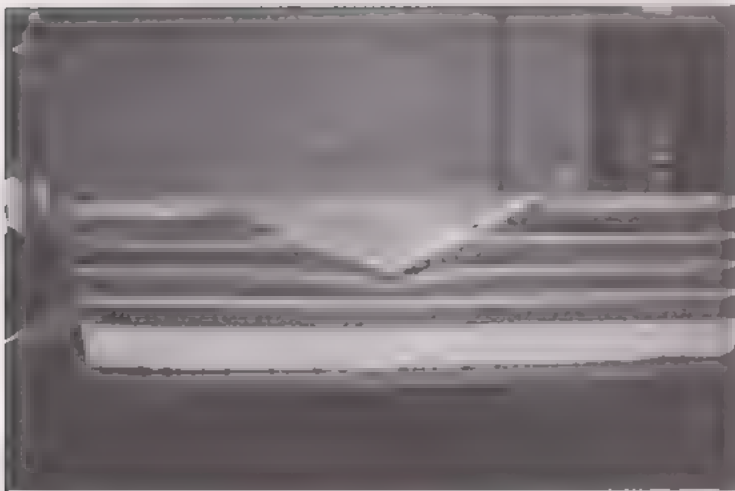
Three lines of evidence show that Meteor Crater was formed by an impact, and not a volcanic eruption. Similar evidence can be used to examine other suspect impact sites.

The 'smoking gun'

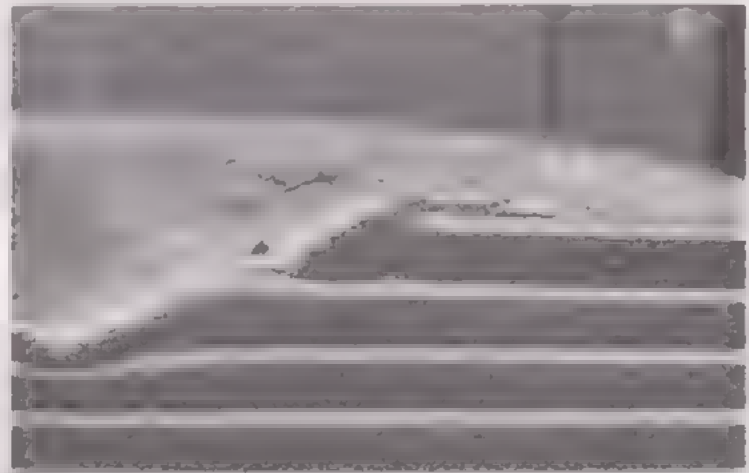
Many large fragments of an iron meteorite have been found on the desert plains surrounding Meteor Crater. Current thinking suggests that the crater was formed by the impact of an iron meteorite less than 50 m in diameter travelling at about 11 km s^{-1} . Many larger impact structures do not preserve fragments of the original projectiles, probably because they vaporized on impact.

Structure

Most volcanic craters are formed by sustained decompression of volcanic gases which blast out large volumes of fragments. Volcanic **ejecta** therefore form simple aprons around craters, the first-erupted materials at the bottom, and the last at the top. Shoemaker's field studies showed that a completely different situation exists at Meteor Crater, where the ejecta form an inverted 'flap', flung out and overturned in a single blast. Thus, a drill hole through the ejecta on the rim would pass through layers of the same (fragmented) material twice. In Shoemaker's words, the strata appeared to 'have been peeled back from the area of the crater, somewhat like petals of a flower blossoming'. Similar structures are produced in laboratory experiments on **hyper-velocity impacts**, those involving closing speeds measured in kilometres per second (Figure 4.3).



(a)



(b)

Figure 4.3 (a) Impact crater made in a NASA laboratory by firing a projectile with velocity of about 1 km s^{-1} into a target of colour banded sand. (b) Detail of the crater rim, showing the overturned 'flap' of ejecta.

Shock phenomena

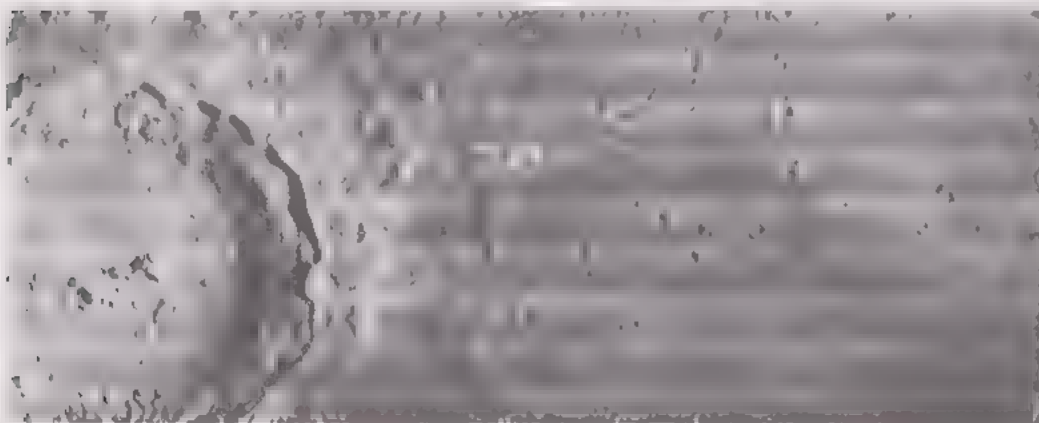
When a projectile hits a solid surface at a velocity of tens of kilometres per second, **shock waves** are set up in the target. A **shock wave** is a sharp compression pulse which radiates in all directions into the target material from the point of impact. Within the compression pulse, pressures may reach levels millions of times greater than the Earth's atmospheric pressure, squeezing rocks so greatly that they may occupy as little as one third of their normal volume. Following immediately behind the compression pulse is a **rarefaction wave**, within which shocked material can flow laterally. Squashing of material around

the impact site forms a **transient cavity**, the first stage in crater formation. As the shock wave advances into the target, shocked material is diverted outwards along the walls of the transient cavity. Decompressed material sprays out of the cavity as an expanding conical curtain of ejecta, often travelling at velocities far greater than that of the original projectile. Ejection of this mass of material constitutes the second, excavation, phase of crater formation. It is all accomplished very quickly – a 1 km crater may be formed in only 10 seconds – but the crater does not take up its final form until the last reverberations of the shock waves have died away, all the ejecta have fallen back, and loose material has avalanched off the steep walls of the cavity.

The enormous transient pressures generated by shock waves leave indelible imprints on the target materials. Conversion of kinetic energy into thermal energy may be sufficient to melt large quantities of rock. At depths of 200–400 m at Meteor Crater, a rubble of shattered target rocks is cemented together by glass, which is formed by shock melting. Pools of comparable solidified **impact melt** line the floors of large lunar craters. Droplets of this melt may be sprayed hundreds of kilometres from the sites of large impacts. Found today in the form of aerodynamically moulded beads of black glass 1–2 cm in diameter, called **tektites**, these melt droplets have been useful in pinning down the sites of several major terrestrial impacts such as that which struck Chicxulub 65 million years ago.

4.2.2 Characteristics of pristine lunar craters

Superb orbital photos of pristine lunar craters (Figures 4.4–4.6 and Plates 2.5 and 2.6) show that the ejecta surrounding them can be divided into several types: a **continuous ejecta** blanket mantles the terrain for a distance equivalent to about one crater diameter, thinning steadily outwards. This continuous blanket forms a hummocky, radially lineated layer of ejecta which completely buries all pre-existing topography, such as older craters. At its margins, patches of underlying terrain begin to show through, defining an area of **discontinuous ejecta**. Within the zone of discontinuous ejecta, numerous **secondary craters** are usually



present, formed where large masses of ejecta, themselves travelling at high velocity, crashed into the surface. It can be difficult to distinguish between secondary craters due to a large impact, and small primary craters. Secondary craters, however, are often clustered together, or aligned in looping chains (Figure 4.6). Swarms of secondary craters also show distinctive 'herring bone' patterns in their own ejecta, radiating away from the primary. This can be helpful in identifying the primary crater that gave rise to the secondaries.

Light-coloured rays of ejecta may extend for thousands of kilometres beyond the obvious ejecta deposits. When the Moon is full, rays from craters such as Tycho present a magnificent sight, as we hope you discovered for yourself in your project work.

Stages of impact crater formation are illustrated in figures in the background information with your Moon Project sheet.

Figure 4.4 A Lunar Orbiter photograph of Copernicus, one of the most closely studied lunar craters. Copernicus (93 km in diameter) is located on the south edge of the Mare Imbrium, and is easily visible with binoculars under good lighting conditions. Morphological points to note at increasing distances outwards from the centre are: (i) a cluster of central peaks; (ii) the flat crater floor, perhaps filled by impact melt; (iii) the multi-terraced crater wall, with terraces formed by inward slumping; (iv) a continuous ejecta mantle, extending out to about 1 crater diameter – features are diffuse within this zone; (v) a discontinuous ejecta apron, with many secondary craters and crater chains; (vi) faint, light-coloured rays extending over the surface of the Mare Imbrium (top). (Lunar Orbiter IV–121H2). This picture is composed of nine strips of film, processed automatically on board the spacecraft, and then scanned and transmitted back to Earth in 1967.

Figure 4.5 Detail of part of the zone of continuous ejecta in Figure 4.4, in an oblique view photographed by an astronaut. Note the subdued hummocky terrain, and pronounced linear features indicating movement from left to right. The hollows are smooth and flat-floored, probably filled by fluidized flows of impact ejecta. The large basin is about 1 km across.

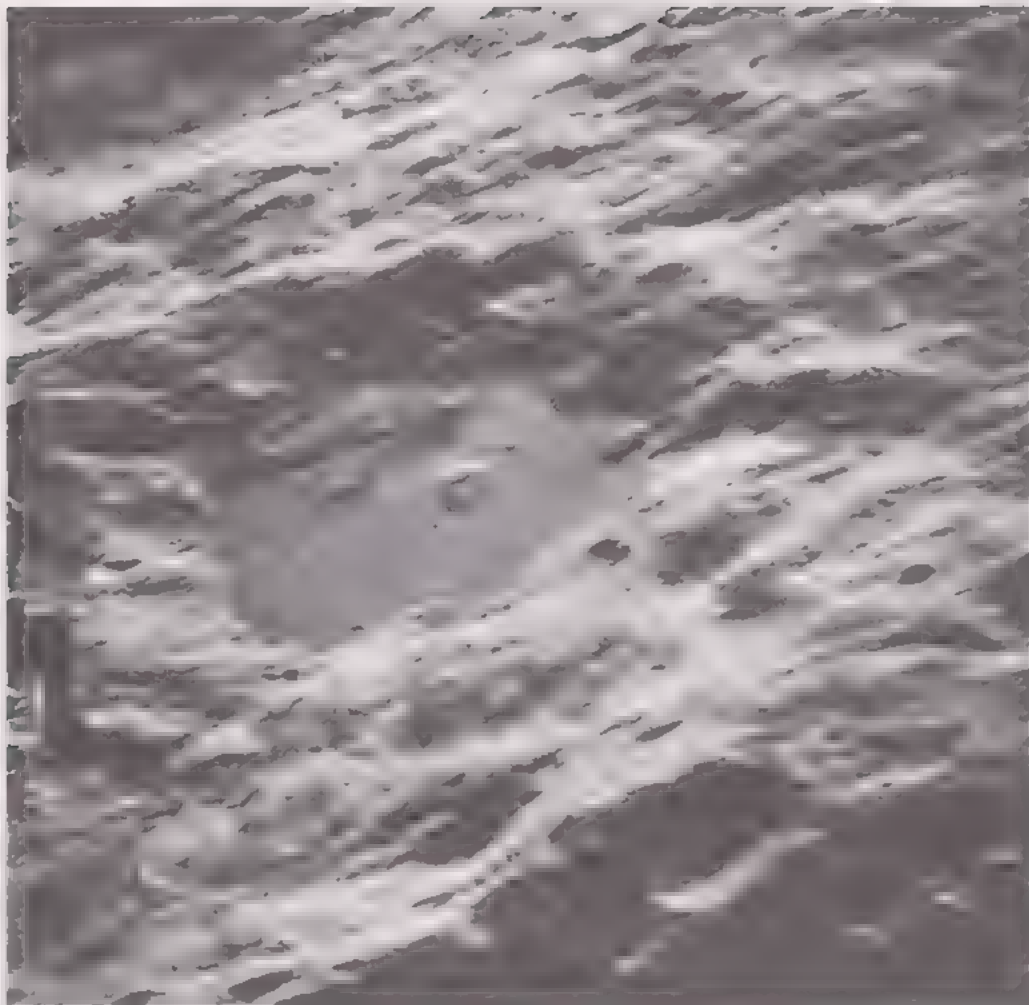


Figure 4.6 An area adjacent to that in Figure 4.4, illustrating chains of secondary craters and 'V'-shaped or herringbone ejecta patterns in the same apron of discontinuous ejecta. Some of the small, perfectly circular craters at the top could be primary. They range in size from 5 km downwards. The image covers an area of about 150 × 170 km. Like Figure 4.4, this is composed of several strips of film; the oblique line of 'bubbles' is a film-processing blemish.



4.3 Shapes of craters

Not all impact craters are perfectly circular. Seen from the air, Meteor Crater has a squarish outline (Figure 4.2). This may reflect the influence of planes of weakness in the target sedimentary rocks.

Can you think of another circumstance, concerned with the trajectory of the projectile, that might yield non-circular craters?

■ If the projectile strikes the target at an oblique angle, an elliptical crater might result.

To a first approximation, projectiles can strike the surface of a planet or satellite at *any* angle. The average impact angle is therefore about 45° .

ITQ 4.1 Examine the photograph of the Moon's surface in Figure 4.7 and estimate quickly the proportion of the craters which are *obviously* elliptical. What inference can you draw from this observation about the process which forms impact craters? (Concentrate on the central part of the photograph – foreshortening will make *all* craters near the edges of the Moon appear elliptical.)



Figure 4.7 An Apollo 17 photograph of the Moon, with craters on its south-eastern flank well illuminated. Most of the area shown lies on the Moon's far side, and consists of heavily cratered highland terrain.

Perhaps contrary to intuition, experiments show that elliptical craters are formed only at angles of incidence that are nearly tangential to the surface. This apparent anomaly arises because hyper-velocity impact craters are formed by essentially instantaneous releases of energy, rather than mechanical distortion. If you throw a pebble at an angle of 45° into a patch of soft mud, it pushes the mud aside, to form an asymmetrical cavity, elongated downrange. If you could hurl the same pebble at a velocity of several kilometres per second at a solid surface, its impact would be *explosive*, producing a circular crater. You could only produce a markedly asymmetrical crater by firing the hyper-velocity projectile at an angle of less than about 5° . When this happens, the projectile partly ploughs through the surface, and the ejecta are thrown outwards on either side of the axis in a 'butterfly' pattern. At very low angles of incidence, long linear grooves may be cut in the surface.

11Q 4.2 Study the photograph of the relatively recent British impact crater in Figure 4.8. What can you say about the likely velocity of the impactor, and the angle of its approach? How do these account for the appearance of the impact site? (*Hint: think about the size of the object, and the likely effects of the atmosphere.*)



Figure 4.8 A suburban setting for one of the smallest terrestrial impact craters. This crater was formed by a meteorite which fell on Christmas Eve 1965 at Barwell, Leicestershire.

4.3.1 Effects of target materials

Many Martian craters look as though they were formed by objects splodging into material with the consistency of wet cement (Figure 4.9). Cratering experts believe that the distinctive ejecta patterns of these **rampart craters** result from the mobilization of trapped ground water, or melting of permafrost ice on impact, yielding ejecta with different properties from those formed on 'dry' planets.

Most of the satellites of the outer planets are composed largely of ice (Subsection 1.2.2, Section 2.6 and Subsection 3.3.2). Intuitively, one might suppose that craters on these bodies should be similar to the 'wet cement' craters on Mars. In fact, they are not – they resemble ordinary lunar craters. One reason

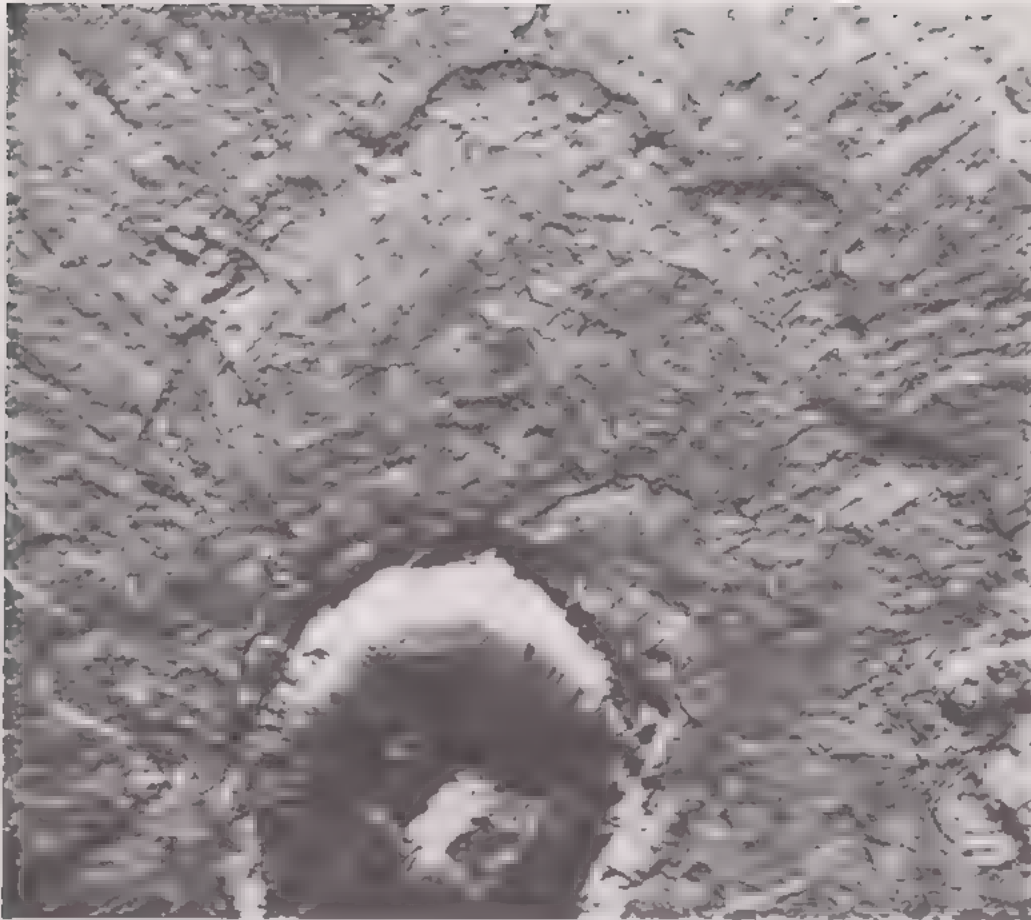


Figure 4.9 Arandas, 28 km in diameter, is a good example of a Martian rampart crater. One prominent lobe of 'flow' material with radial striations extends out to about 1 crater diameter. Two or three other series of lobes are present, extending beyond the edge of the image. One lobe reaches more than 50 km from the crater rim, much further than one would expect for continuous ejecta that travelled purely ballistically. Striations appear to extend inwards as far as the central peak.

for this may be that surface temperatures on the satellites are as much as 200 K lower than on Mars. At these frigid temperatures, ice is mechanically much stronger than at higher temperatures and resembles low-density rock.

ITQ 4.3 Thick masses of ice on Earth display a property which radically affects their shape over time. What is this property, and how might it affect the shapes of craters on icy satellites?



Figure 4.10 A Voyager spacecraft picture of the vast Valhalla impact site on Callisto. Valhalla is the largest known impact structure in the Solar System, with a diameter of 4 000 km, yet its topography is so subdued that each of the numerous visible ring scarps rises only 1–2 km above its surroundings. The impact which formed Valhalla must have taken place at a time when Callisto's lithosphere was thin, and its asthenosphere hot and mobile enough to smooth out the surface topography by viscous relaxation. The whitish zone near the middle is the palimpsest on the site of the (now vanished) transient cavity.

The word *palimpsest* originally referred to a manuscript in which earlier writings had been rubbed out, so that it could be used again.

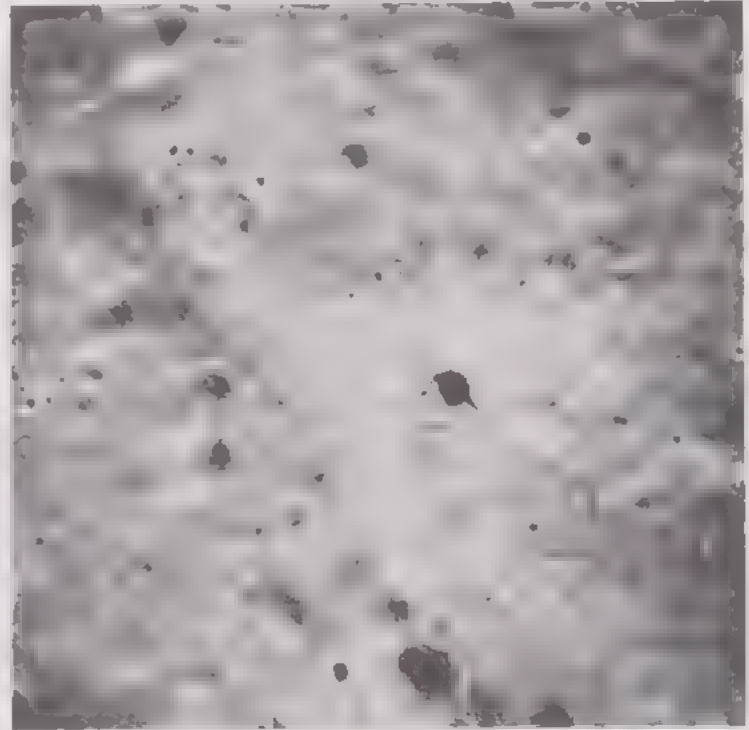
Because normal surface temperatures are today so low on planetary satellites, ice near the surface is mechanically much stronger than on Earth, forming a lithosphere in its own right as we saw in Subsection 3.5.3, so the possibilities of glacier-like plastic flow are limited. Early in the history of the Solar System, surface temperatures of icy satellites may have been higher than they are today, permitting some glacier-like flow of surface ice structures to take place. Furthermore, the icy lithospheres of large satellites would inevitably have been thinner early in their history, when their interiors were warmer. Thus, surface topographic effects would become subdued as the volume of asthenosphere that was deformed on impact gradually returned to its original position by **viscous relaxation**. This, and glacier-like flow, has led to the formation of crater **palimpsests**, which are ghostly imprints of former craters. Examples of these are visible on Jupiter's large icy satellites, Ganymede and Callisto (Figure 4.10). Viscous relaxation also affected rocky bodies such as the Moon, when very large impacts deformed both its rigid lithosphere, and its plastic asthenosphere.

4.3.2 Effects of scale

Figure 4.11 illustrates some lunar craters of different sizes ranging from a few tens of metres to several hundred kilometres. Look carefully at these craters, and compare them with one another.

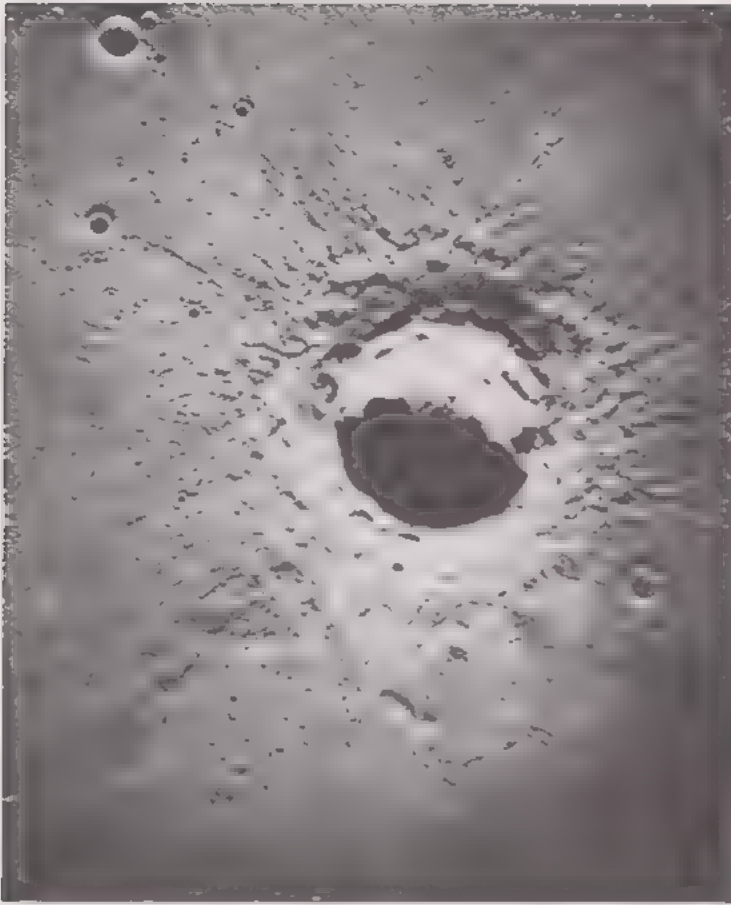


(a)

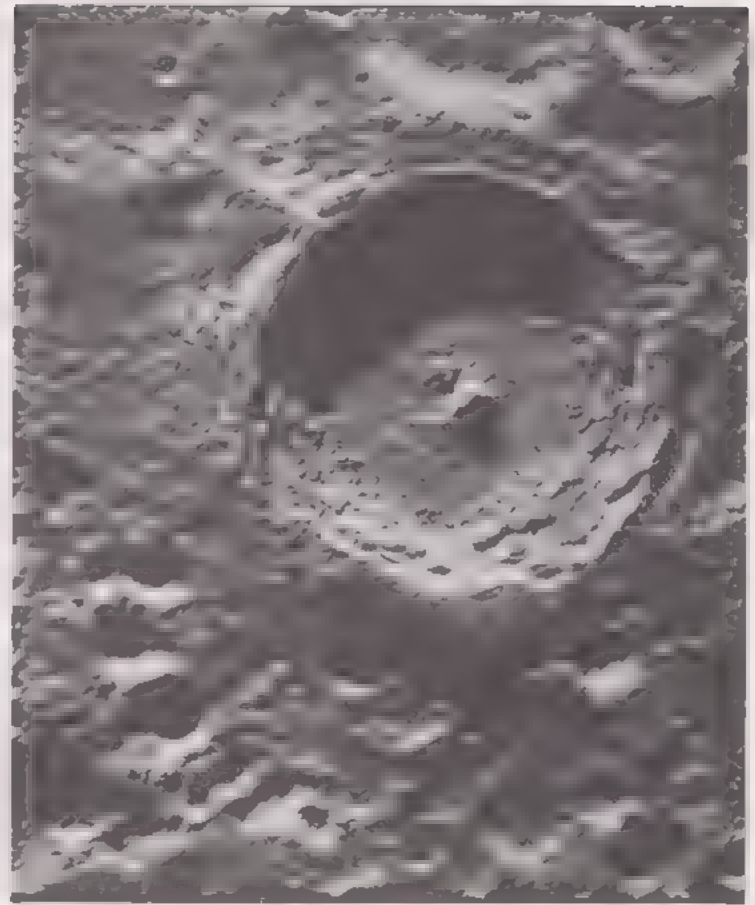


(b)

Figure 4.11 (part) Variations in structure of lunar craters with scale. Differences in lighting and viewing geometry complicate comparisons. (a) An unnamed far-side crater about 3 km across, photographed from the Apollo 10 lunar module. It is a simple, deep bowl-like shape, with materials slumped from the walls. (b) A young unnamed far-side crater about 10 km in diameter, with prominent radial rays of light-toned ejecta.



(c)



(d)



(e)

Figure 4.11 (contd.) (c) Timocharis crater, 32 km across, on the Mare Imbrium. Well-developed continuous and discontinuous ejecta mantles are present. There is a central peak (shadowed), and simple terracing of the crater wall – one or two benches can be seen. (d) Tycho, the youngest major crater on the Moon, is about 80 km across. The crater floor is flat and expansive; there is a compound central peak, and multiple benches are present on the crater walls. Tycho is thought to be about 100 million years old. (e) The Orientale basin on the lunar far side, 900 km in diameter. Major concentric ridges are present, but there is no central peak complex. There are some small patches of lava in the outer ring, and the centre of the basin has been flooded by lavas.

Are the largest craters merely scaled-up versions of the smaller ones, or are there systematic differences?

■ Large craters are markedly more complex than smaller ones. In particular, the largest impact structure, the Orientale Basin (900 km across), exhibits a conspicuous ‘bull’s eye’ pattern of concentric rings.

Statistical analyses show that as craters get bigger, there is a progressive decrease in the ratio of their depth to diameter, so that while small craters resemble deep bowls, large craters are more like shallow dinner plates. A loose hierarchy of lunar crater structures can be seen as follows:

Craters less than 1 km in diameter are simple pits with bouldery ejecta aprons.

Craters between 1 and 15 km in diameter are deep, bowl-shaped depressions.

Craters between 15 and 140 km in diameter are shallower, have flat floors and are more complex, showing evidence of slumping from crater walls to form **terraces** and may also have single **central peaks**.

Craters between 140 and 175 km in diameter tend to have clusters of central peaks, or a fragmentary central ring.

Craters more than 175 km in diameter lack central peaks, but possess more than one ring and are termed **peak-ring basins**.

Craters more than 350 km in diameter are complex, concentric multi-ring structures, often termed **multi-ringed basins**, such as the Orientale basin on the Moon (Figure 4.11e).

Because the effects of large impacts were so profound, effectively puncturing the lunar lithosphere, lavas often welled up into impact basins and large craters, masking many original features. The dark lunar **maria** (pronounced ‘mah-ria’; singular, **mare**, ‘mah-ray’ – from the Latin for sea) are all impact basins flooded by lavas. Radiometric dates (Box 3.1) of samples collected on the Moon suggest that this volcanic flooding continued for hundreds of millions of years after the impacts that excavated each basin.

Most of the attributes of craters of intermediate size can be explained in terms of the effects of impact shock waves on the lunar lithosphere. Central peaks, for example, in craters more than 15 km in diameter are due to the central region of the impact ‘bouncing back’. Something similar happens when water drips into a bowl: when a drop *first* hits the surface, it depresses the surface and produces an outwards-propagating wave ring, but then the central region jumps visibly upwards, sometimes ejecting a tongue of liquid (video sequence 6). High-speed photography conveniently freezes this spurt; on the Moon, the rapid dissipation of shock-wave energy left central peaks prominent, frozen in time.

Further complications result if the size of the crater is the same order of magnitude as the thickness of the lithosphere, or greater. In these cases, shock waves penetrate both the rigid lithosphere and the plastic asthenosphere, behaving differently in each. Interactions between the two are responsible for formation of complex multi-ringed basins.

Because the same physical processes operate throughout the Solar System, similar changes in crater structure with size can be found on any solid object. Large craters on Earth, such as Manicouagan, for example, show beautiful (though much eroded) concentric structures similar to those on the Moon.

Would you expect to see *exactly* the same variations with size of crater on other solid bodies? If not, what factors might influence shapes of craters?

■ Three factors which influence the shapes of craters of varying sizes are the surface gravity of the body; the thickness of its lithosphere; and the strength of the materials of which it is made.

If impacting bodies have the same masses and arrive at the same velocity, they will excavate smaller craters on planetary bodies with high surface gravity than on low-gravity bodies (Table 4.1). Less obviously, on high-gravity bodies, central peaks and rings would also appear in craters of smaller diameters. Thus, central peaks are seen in terrestrial craters as small as 6 km in diameter, compared with 15 km on the Moon. When craters reach dimensions of hundreds of kilometres and have complex concentric structures, it is less easy to ascribe variations to individual parameters such as surface gravity or lithospheric thickness.

Table 4.1 Diameters of craters on the terrestrial planets that would be equivalent to a 10 km crater on the Moon

Planet	Crater diameter / km	Surface gravity ^a / m s ⁻²
Mercury	7.2	3.7
Venus	5.1	8.9
Earth	4.9	9.8
Moon	10.0	1.6
Mars	7.2	3.7

^a This is the acceleration due to gravity at the surface of the planet (*Preparatory science*, Subsection 1.4.1).

Summary of Section 4.2 and 4.3 and SAQs

- 1 Impact craters have played a major role in the histories of all the solid objects in the Solar System, including the Earth. Though superficially similar, they may be distinguished from volcanic craters through evidence of shock phenomena and by overturned flaps of ejecta.
- 2 Pristine craters exhibit distinctive patterns of continuous and discontinuous ejecta. Thin 'rays' of ejecta from recent impacts extend for thousands of kilometres on the Moon.
- 3 Elliptical craters are formed only by oblique, low angle impacts ($< 5^\circ$); other craters are circular, but show a range of increasingly complex structures with increasing size. Small lunar craters (1–2 km in diameter) are simple, deep bowls; with increasing size, craters become relatively shallow and have flat floors.
- 4 Lunar craters more than about 15 km in diameter have more complex structures, including central peaks, while those larger than about 175 km are concentric ring structures. Crater sizes and morphology are also influenced by the surface gravity of the target body, the nature of the surface material, and the thickness of the lithosphere.
- 5 On Mars, 'rampart' craters suggest that impacts took place on a surface containing permafrost ice.
- 6 On satellites with relatively thin lithospheres, large impacts may cause deformation of the underlying asthenosphere; when this recovers by viscous relaxation the surface structure may be partially erased.
- 7 On icy satellites, very cold ice has similar strength to that of low-density rock, but may exhibit glacier-like flow at higher temperatures.

SAQ 4.1 (Objectives 4.2–4.4 and 4.9) Look closely at Figure 4.12. What can you say about the origin of the small craters in the photograph? One major crater, Copernicus, is visible at the top right. The photo also contains evidence for the existence of a second large crater. In which direction does it lie?

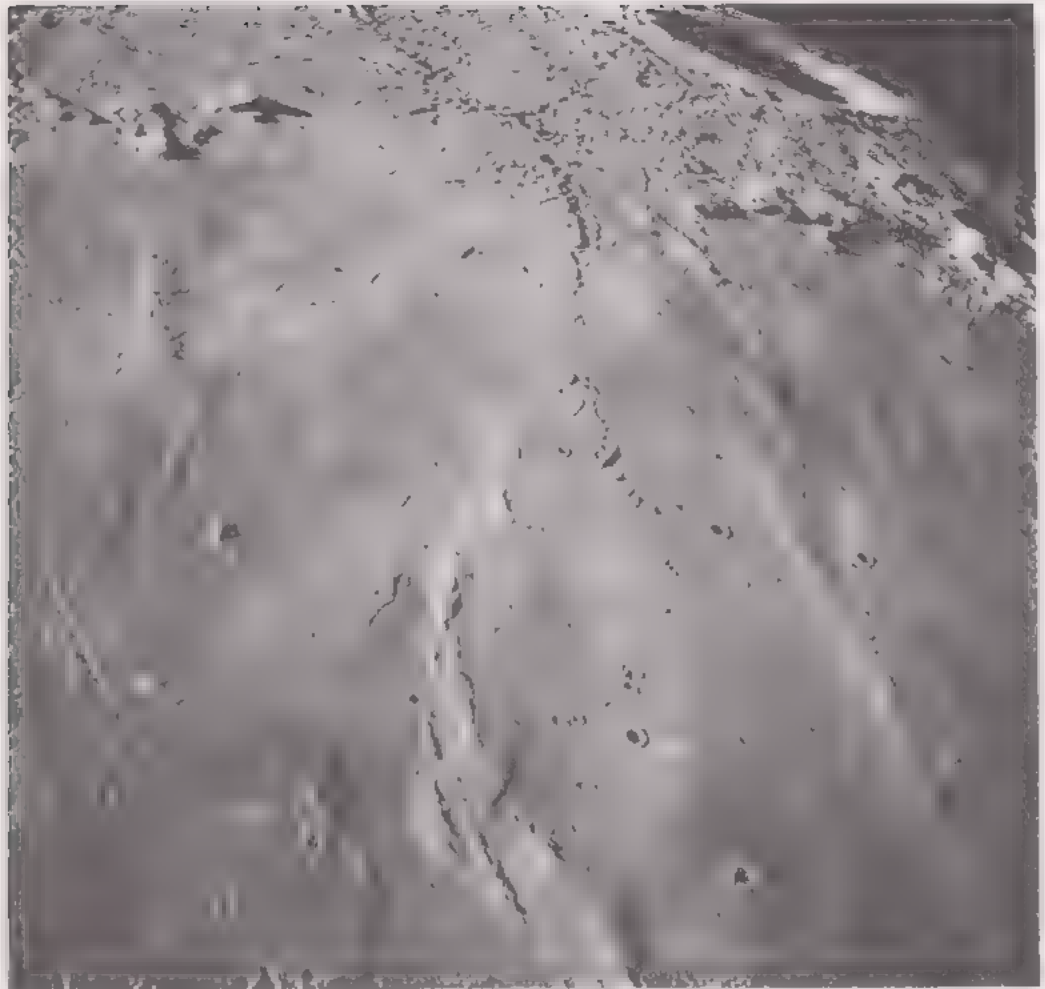


Figure 4.12 An Apollo photograph of Copernicus crater (top right), for use with SAQ 4.1.



Figure 4.13 An Apollo photograph of the Moon's surface in the Oceanus Procellarum, for use with SAQ 4.3.

SAQ 4.2 (Objectives 4.4 and 4.9) The 900 km diameter Orientale basin on the Moon shows less evidence of viscous relaxation than the Valhalla basin on Callisto – its rings are far more prominent topographic features. Account for this difference in terms of what you know about the origin and thermal evolution of planetary bodies.

SAQ 4.3 (Objectives 4.2–4.4) Examine Figure 4.13. Ignoring the large ring-like structure, what can you say about the origin and maximum likely size of the craters in this view?

4.4 Craters as chronometers

In the lead-up to the Apollo program, each refinement in technology revealed more craters, until images from the first spacecraft to land revealed that the surface is cratered down to a scale of millimetres. Lunar scientists recognized that the numbers of craters on a surface could be used to estimate its age: the longer a surface has been exposed to impacts, the more craters it will exhibit. Although simply stated, however, the technique is far from simple to apply.

Before reading on, jot down some of the factors that you consider may cause complications.

■ Five important factors are:

- (i) The rate at which impacting bodies strike the surface
- (ii) The size distribution of impacting bodies
- (iii) The obliteration of early-formed craters by later ones
- (iv) The effects of secondary impacts
- (v) The need to establish an *absolute* frame of reference.

We shall consider each of (i)–(v) in turn.

(i) *The rate at which impacting bodies strike the surface*

The number of impacting bodies hitting unit area in unit time is termed the **impact flux**. If the rate at which a surface is being bombarded varies through time, then the cratering chronometer becomes harder to use. In the wake of the formation of the solid planets there was a period of intense bombardment, but this waned rapidly until the flux declined to something like its present level. This pattern seems to have been consistent within the inner Solar System, so crater statistics do provide useful chronological clues.

(ii) *The size distribution of impacting bodies*

Photographs of the Moon show that for every large crater there are many smaller ones. Thus, if we are to use crater statistics to estimate ages, we have to decide how to manage the variations in size of craters. Furthermore, the flux of impacting bodies may not always have had the same **size-frequency distribution**. Size-frequency distribution is a statistician's term used to describe the relative numbers of objects across a range of sizes. In a **normally distributed** population, there will be a few very small objects and a few very large objects, while most will cluster around some intermediate size.

□ Can you name a variable in the human population that shows normal distribution?

■ Adults' heights are normally distributed – there are few very short or very tall people, but many around 'average' height. (Adults' weights are an equally good example.)

The statistics for men and women show very similar distribution patterns, but the 'average' for men is slightly greater than that for women.

In cratering, the 'usual' flux is *not* normally distributed – there are far more small bodies than large ones, but there may also be times when a surface is hit by showers of projectiles of closely similar size, like shotgun pellets, and other times when a mixture of projectiles of a much broader range of sizes bombards the surface. Thus, varying size frequency distributions can lead us into some difficult statistics. But one conclusion is clear: it is essential to understand the impactors before one can understand the impact craters.

Can you suggest an imaginary experiment that would throw light on both the current impact flux on the Moon, and the size-frequency distribution of the craters that result?

■ Ideally, one would scrape a surface clear, wait a few million years, and then count the number of craters formed within each of a number of size intervals.

In effect, this experiment has already been done for us over a longer time interval – an important discovery from radiometric dating (Box 3.1) of lunar samples returned by the Apollo missions was that the mare basins were completely covered by lavas more than 3 billion years ago. They still retain many features of the floods of lava, which effectively wiped away all records of earlier impacts. Thus, measurements of the size-frequency distribution of craters on mare surfaces yield the **crater production flux** over the last three billion years or so, and the size-frequency distribution of the resulting craters. The surfaces of the maria range in age from about 3.9 to 3.2 billion years; the 'average' mare age is about 3.6 billion years.

(iii) The effects of obliteration

If projectiles are continuously fired at a surface, early-formed craters will eventually be obliterated by younger ones. When a new crater can form only at the expense of an older one, by overprinting it, the surface is said to be **saturated**. Clearly, the pre-saturation history of a saturated surface is irretrievably lost.

IIQ 4.4 Examine the two pictures in Figure 4.14. Does either of them reveal a saturated surface? What can you say about the relative ages of the two surfaces illustrated?

(iv) The effects of secondary impacts

Crater-counting statistics become unreliable if the origin of the craters is dubious. Large impacts produce huge numbers of secondary craters, so problems can arise if these are included within the population of primary impacts. Statistical techniques have been devised to cope with this issue, such as not counting small craters, many of which may be secondaries.

(v) The need for an absolute frame of reference

Crater counting by itself offers, at best, only a means of comparing *relative* ages of surfaces. But because samples returned from the Moon have been dated by radiometric techniques, the Moon provides a calibrated measuring stick, enabling us to use crater statistics in an *absolute* time frame. Comparative crater statistics can be used to estimate reliably the ages of lunar surfaces which have not been dated in the laboratory. But this confidence does not extend far. In the absence of any other reference points, we *also* have to use the lunar calibration to estimate the ages of surfaces on Mercury, Venus and Mars. There are reasons, however, to suppose that the impact flux in the *outer* Solar System was significantly different, both in its size distribution, and through time. Thus, we have only rather poor absolute estimates of the surface ages of most of the solid bodies in the Solar System.

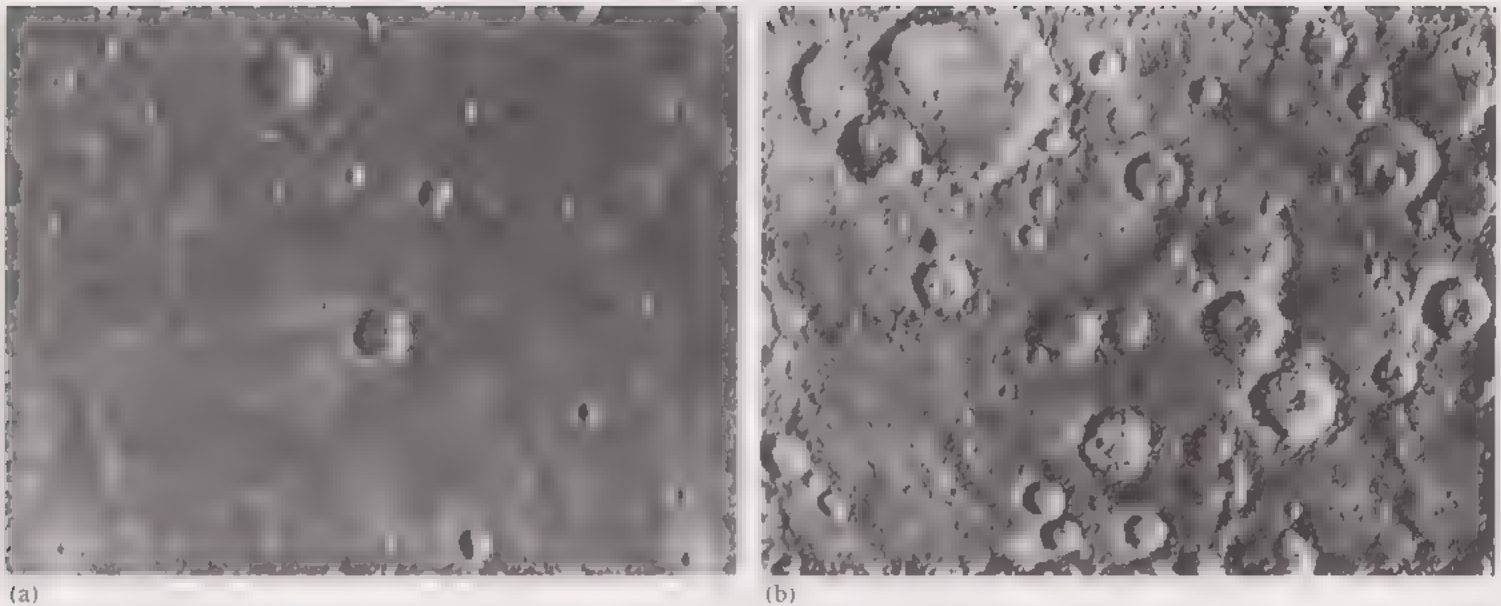


Figure 4.14 Pictures of planetary surfaces, for use with ITQ 4.4.

4.4.1 Plotting crater size frequency distribution curves

Crater statistics are usually displayed on graphs of crater frequency against diameter, similar to that in Figure 4.15. There are two complications involved in plotting and interpreting these graphs.

First, although the horizontal axis, *crater diameter*, is logarithmic, each interval represents a doubling in crater diameter, rather than a power of ten.

Why use a logarithmic scale at all? Why not plot crater diameters on a linear scale?

■ The range in diameters is so great, from 1 to 512 km, that a linear graph would be ridiculously long. Furthermore, whereas an increment of 1 km from 2 km to 3 km represents a major *relative* change, an increment of 1 km from 100 km to 101 km is a much smaller relative change.

The vertical axis of the graph ('powers of ten' logarithmic) is not simply 'number of craters per square kilometre', but the number of craters falling *within a given size range* per square kilometre of the surface. Numbers within incremental sizes are used because they reveal information about the cratering process. Choosing a doubling scale for the horizontal axis – powers of 2 rather than 10 – provides a more convenient scale, with more 'bins', or divisions, in the smaller crater size range.

Working through an example should clarify most of these issues. Table 4.2 lists actual crater statistics for the Earth, in terms of actual numbers, and numbers per square kilometre.

Table 4.2 Numbers of craters in given size ranges found in ancient, undeformed areas of the Earth's continents

Crater size range / km	Number of known craters falling within the size range	Number of craters in the size range per km ²
1–2	7	4.7×10^{-8}
2–4	12	8.0×10^{-8}
4–8	15	1.0×10^{-7}
8–16	17	1.1×10^{-7}
16–32	16	1.1×10^{-7}
32–64	5	3.3×10^{-8}
64–128	3	2.0×10^{-8}
128–256	2	1.3×10^{-8}

ITQ 4.5 Plot these data on the blank graph in Figure 4.15. Note carefully how the divisions are arranged. It is easy, for example, to suppose that 3×10^{-8} is a larger number than 3×10^{-7} , whereas the reverse is in fact true. Sketch a curve through the points you have plotted. What conclusions can you draw from the shape of your curve about the frequency of small craters on Earth?

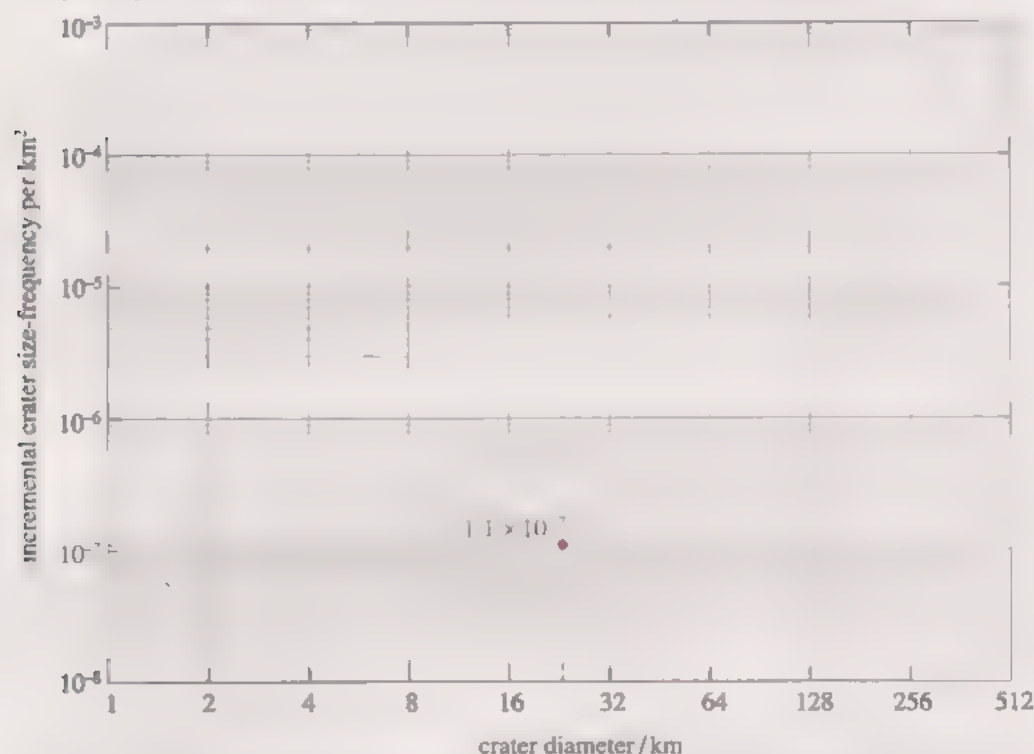


Figure 4.15 Graph for plotting crater size-frequency statistics (ITQ 4.5). The horizontal axis represents crater diameter in kilometres. Note that it is logarithmic, each unit is double the previous one. The vertical axis, also logarithmic, is incremental size-frequency per square kilometre, i.e. the number of craters within each size 'bin' per square kilometre. To get you started, the 16–32 km crater diameter data point is plotted as an example.

The irregularity of the curve in Figure 4.22, and the paucity of small craters, result from erosion taking place on the Earth's surface, which erases small craters more easily than large ones. Almost all the craters plotted are less than 500 million years old, and only one is more than 1 000 million years old. Figure 4.16 shows comparable crater size-frequency distribution for lunar mare surfaces, which average about 3.6 billion years old.

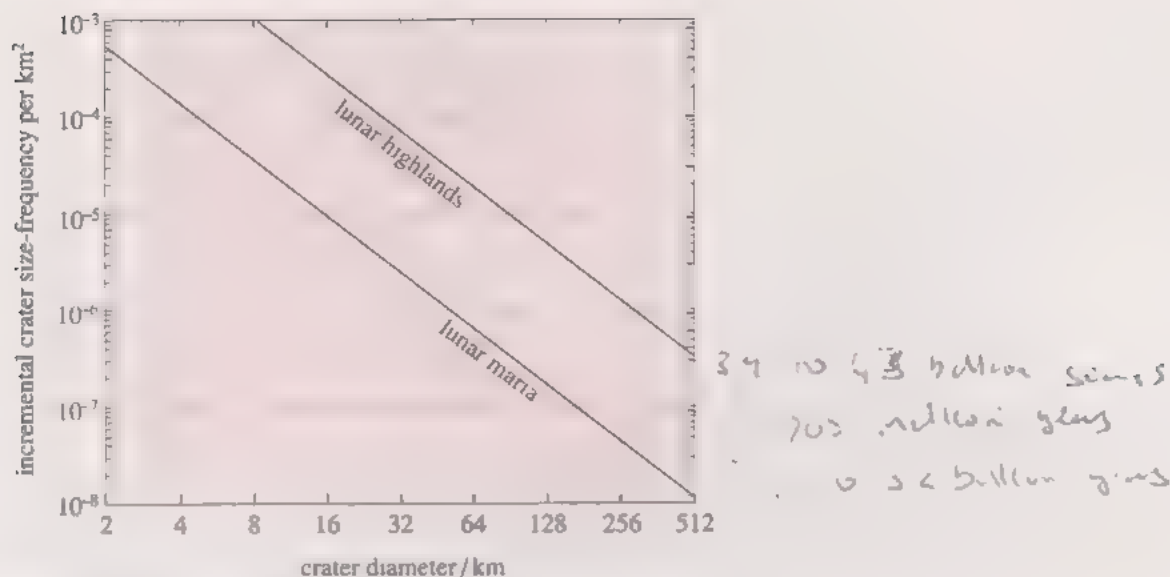


Figure 4.16 Crater size-frequency distributions for the lunar maria and highlands.

Pencil the curve from Figure 4.22 onto Figure 4.16 and compare it with the curve for the lunar maria. What is the difference in crater frequency for craters 16–32 km in diameter between the two curves? How can one account for these differences?

■ There are about 10^{-7} craters 16–32 km in diameter per square kilometre on the Earth, but about 5×10^{-6} on the maria, almost two orders of magnitude, or 100 times more. The difference reflects two factors: the long exposure of the Moon's surface to space, and the far slower rates of geological processes there which can erase craters once formed.

The topmost curve in Figure 4.16 shows crater statistics for the lunar highlands, which are about 3.9–4.3 billion years old. These surfaces are saturated: new craters can only be formed at the expense of older ones.

Where would you expect the cratering curve for a lunar surface 4.5 billion years old to plot?

■ Any surface older than 4.3 billion years must also inevitably be saturated. Thus such a surface would also plot along the 4.3 billion years line, and the space above it would be 'forbidden territory'.

We now have cratering curves for two distinct terrain types on the Moon, both of which have been radiometrically dated. These established curves enable us to begin to use cratering as a chronometer, and say something about the mixture of sizes of bodies impacting on the Moon.

4.4.2 Using craters as chronometers

The differences in crater statistics for the three terrains in Figure 4.16 (lunar highlands, lunar maria and your curve for Earth's surface) permit one to estimate the production rate for craters of a given diameter. Considering craters 32–64 km in diameter, there are about $4 \times 10^{-5} \text{ km}^{-2}$ on the highlands, and $1 \times 10^{-6} \text{ km}^{-2}$ on the maria. Given that the age difference between the two surfaces is 700 million years, this shows that 3.9×10^{-5} craters 32–64 km in diameter were formed per square kilometre in the 700 million years prior to 3.2 billion years ago – a rate of $5.6 \times 10^{-5} \text{ km}^{-2}$ per billion years.

ITQ 4.6 To examine the reality of that rather arcane statistic, consider the Mare Serenitatis, which has an area of about 200 000 km² and whose surface has been dated from returned samples at 3.2 billion years old. How many 32–64 km craters would have been formed from the time of origin of the mare surface to the present day, if the production rate of 5.6×10^{-5} km⁻² per billion years was sustained?

Actual counts of the numbers of 32–64 km craters on Mare Serenitatis – which you could confirm with binoculars – shows that there are hardly any craters of this size. This discrepancy arises because the impact flux did not remain constant from 3.9 billion years ago (the age of the last heavy bombardment) to the present. In fact, detailed evidence indicates that it declined steeply initially and more slowly later.

ITQ 4.7 Figure 4.16 shows that the frequency of craters 32–64 km in diameter on the lunar maria is about 1.5×10^{-6} km⁻². There are about 2 craters within the same size range on the Canadian Shield, which has an area of about 10⁷ km², and is one of the best studied areas of old, little-deformed rocks on Earth. If the crater production flux represented by the lunar maria (i.e. over the past 3.2 billion years) could be *directly* applied to the Earth, how old would you expect the Canadian Shield to be?

Given that the oldest independently dated craters on the Canadian Shield are about 500 million years old, your answer to ITQ 4.7 should show that crater counting provides a reasonably good estimate of the age of its surface, and demonstrates the usefulness of crater statistics in making planetary comparisons. However, there are some further complications that we need to explore.

4.4.3 A calibrated cratering curve

One of the striking things about Figure 4.16 is that the cratering curves for the lunar highlands and lunar maria have the same gradient, suggesting that the impacting bodies had similar size-frequency distributions. Sophisticated statistical manipulations reveal some subtle differences between the two populations, and that many of the myriads of small craters (less than 4 km in diameter) on the mare surface may be secondaries. Compilation of crater statistics from all radiometrically dated areas of the Moon has enabled the plot in Figure 4.17 to be constructed. This graph is the key to using craters as chronometers in the Solar System. Because the lunar mare surfaces are best dated, crater densities for other planetary surfaces are 'normalized' to the mare value, i.e. expressed as fractions of the mare value. Also, rather than utilizing simply the number of craters in each or any size bin, the *total* number of all craters greater than 4 km diameter is used.

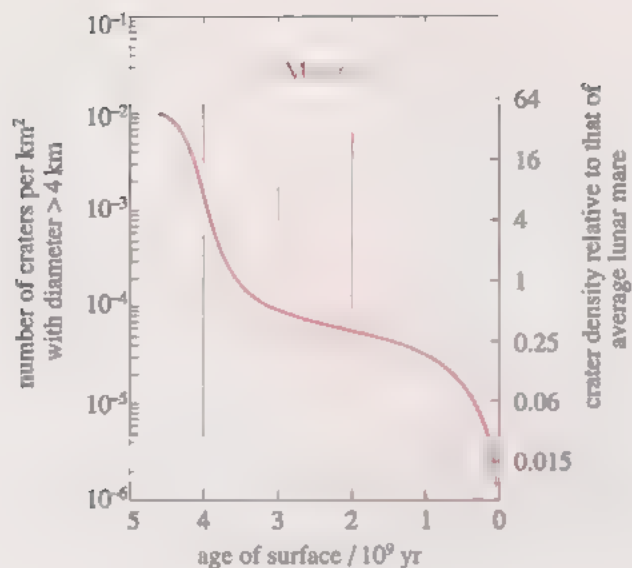


Figure 4.17 'Calibrated' cratering curve for the Moon, based on statistics for surfaces whose ages have been determined radiometrically. The left vertical axis shows number of craters per square kilometre with diameters greater than 4 km. (Craters smaller than 4 km may be secondaries.) The right vertical axis shows crater densities as multiples of the crater density for the average lunar mare, the best dated type of surface. The shaded area represents the range of errors in dating and crater counting.

4.4.4 Comparisons with Mars

Because no surfaces have yet been sampled from Mars for radiometric dating, crater statistics have been minutely examined to estimate the age of surface features. Table 4.3 lists a comparison between some Martian surface features and the lunar maria.

Table 4.3 Crater densities of some Martian geological terrains (ages to be added in ITQ 4.8)

Geological province	Crater density relative to that of average lunar mare	Estimated age / 10^9 yr
Tharsis volcanic plains	0.1	1
Elysium volcanic plains	0.7	2
Chryse Planitia	1.1	3
Tyrrenum Patera	1.4	3½
Hellas	1.8	3½
Syrts Major	2.0	3½
Heavily cratered plains	13.0	4

Excellent images of Mars are available, so these crater *numbers* are themselves secure, but before we can use them as a chronometer for Martian geological processes, we need to understand how the cratering process on Mars compares with that on the Moon. Because Mars has a larger mass than the Moon, gravitational focusing (Subsection 2.5.1) will attract more impactors towards it, thus generating a higher flux. Mars is also much closer to the orbiting 'scrapyard' of the asteroid belt, so there are more potential impactors.

Both these factors indicate that Mars experienced a larger impact flux than the Moon. Exactly how much larger is controversial, but the current 'best estimate' is that it was roughly a factor of 2 larger.

Figure 4.18 is basically the same plot as Figure 4.17, shifted vertically to take account of the higher overall flux at Mars. Similar figures can be constructed for the other planets and satellites, with varying degrees of confidence.

Remember that impact flux is quoted *per unit area* (Section 4.4).

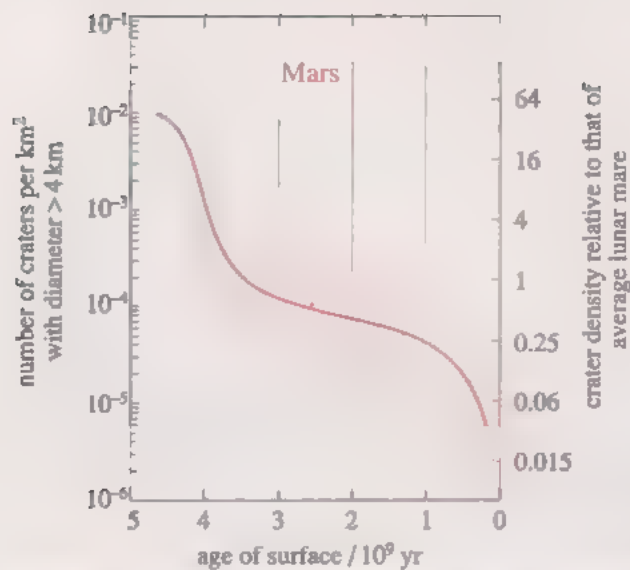


Figure 4.18 Cratering curve for Mars, based on lunar data in Figure 4.17 but adjusted to take account of the higher flux at Mars.

ITQ 4.8 Use Figure 4.18 and the data in Table 4.3, to estimate ages for the Martian surfaces listed, writing your estimates in the third column. (Given the scale of the graph, and the magnitude of the uncertainties in the data, you will only be able to make rough estimates.) What do your answers reveal about the geological history of Mars compared with that of the Moon?

As your answer to ITQ 4.8 should show, Mars's *youngest* features (the Tharsis volcanoes) are still ancient by comparison with those of the Earth, where, for example, 300 million years ago the Atlantic Ocean did not exist. Over the last 2–3 billion years, Mars's cratering record has been similar to that of the lunar maria. What makes its earlier record complex is that an episode of rapid erosion appears to have taken place while the waning stages of heavy bombardment were still continuing (Figure 4.19). Although Mars exhibits abundant evidence of erosional processes, there appears to have been relatively little *recent* erosion over large areas.

In the next chapter we shall examine the processes that have contributed to the formation of Mars's Tharsis volcanoes, and to volcanism on planetary bodies in general.

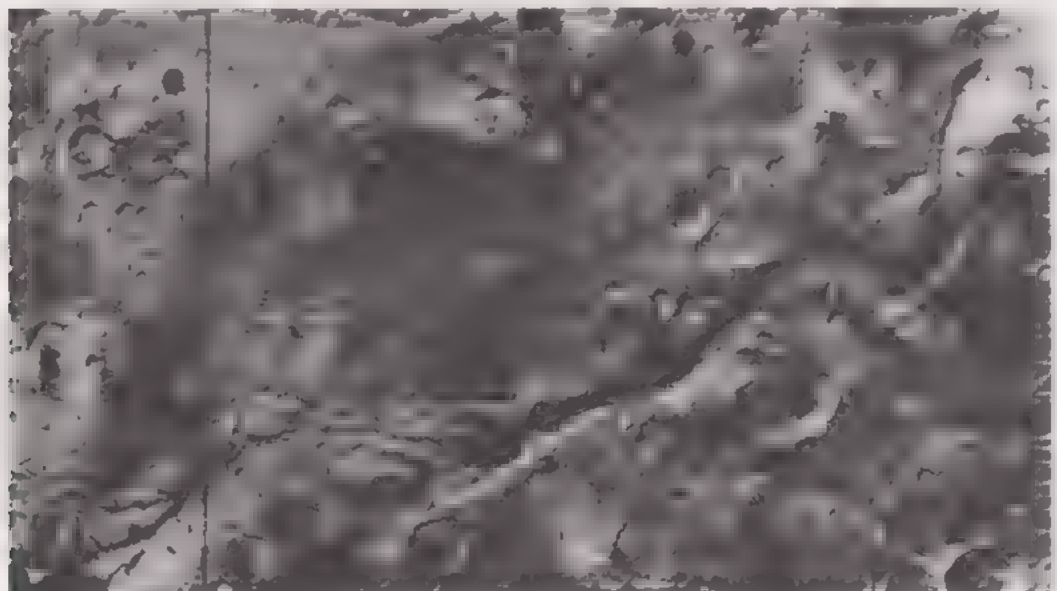


Figure 4.19 Part of a 700 km long channel in the heavily cratered southern highland region of Mars discovered by Mariner 9 in 1972. Channels like these provide firm evidence of an episode of erosion by flowing water, very early in Mars's history, before heavy bombardment had ceased entirely. Liquid water could not exist on Mars at the present day.

ITQ 4.9 How would you expect Mars's early episode of erosion to be manifested in the distribution of crater sizes for the oldest terrains? Try and sketch roughly what you think the incremental size-frequency curve should look like on a plot such as that in Figure 4.16.

Summary of Section 4.4 and SAQs

- 1 Crater size-frequency distribution statistics are invaluable because (in the absence of radiometric dating) they provide the only objective means of comparing the relative ages of surfaces of solid planetary bodies within the Solar System.
- 2 Variations in the impact flux through time, and its size-frequency distribution, coupled with the effects of overprinting of older craters by younger ones, complicate statistical analyses.
- 3 Secondary impacts yield anomalously large numbers of small craters, while erosion on planets such as Earth and Mars may skew the distribution statistics by erasing small craters.
- 4 Radiometric dating of surfaces on the Moon provides a means of calibrating crater size-frequency distribution statistics, and this calibration can be extended to other planets, if 'corrections' to the assumed flux can be made to account for variations in factors such as planetary gravity, and location within the Solar System.

SAQ 4.4 (Objectives 4.4–4.6) Examine Figure 4.20, which is an Apollo photograph of the lunar surface, about 25 km × 30 km across. (a) Look first at the area to the left of centre. How would you explain the crowded clusters of craters that are visible? (b) Are there obvious differences in crater size-frequency distributions in different parts of the picture? If so, how would you explain them? (c) Assuming that you devoted sufficient time to counting the craters, discuss whether you could make meaningful estimates of the age differences between different areas.

SAQ 4.5 (Objectives 4.6–4.8) Figure 4.21 shows cratering statistics derived for the Earth, similar to those for the Moon and Mars in Figures 4.17 and 4.18. Use the data in this figure to estimate the crater production rate on the Earth *relative* to that on the Moon. Comment on the relative rate, and the implications of the uncertainties in the data for the reliability of the value you have obtained.

SAQ 4.6 (Objectives 4.5–4.7) Much of the discussion in this section has concerned crater size-frequency distributions with similar gradients, reflecting impacts by objects with similar size-frequency distributions. Describe the appearance of the crater size-frequency distribution plot for a planet which had been bombarded by impactors whose size-frequency distribution changed with time, so that initially it was dominated by mostly very small objects, and subsequently by impactors with the usual size-frequency distribution. Assume that bombardment continued to the point of saturation.

SAQ 4.7 (Objectives 4.5–4.7) Examine the craters in Figure 4.14b, which is a Mariner 10 image of the surface of Mercury. What can you say about the size range in which central peaks are present? (Rabelais, at the top left of the picture, is 120 km across.) How does this compare with the Moon?



Figure 4.20 Apollo photo of the lunar surface, for use with for SAQ 4.4.

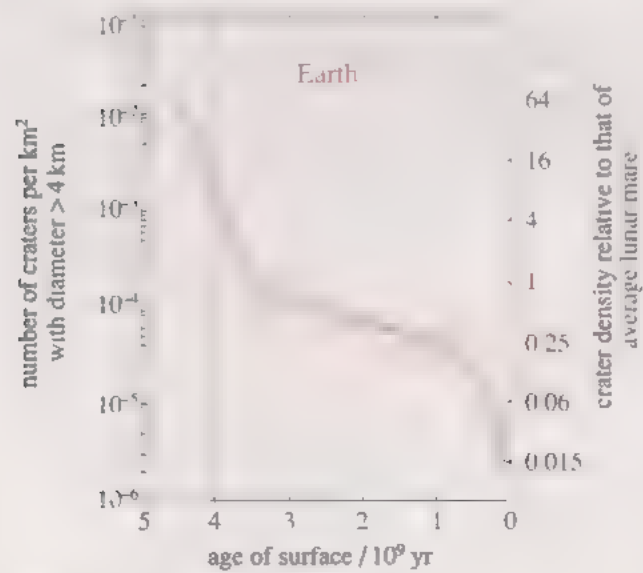


Figure 4.21 Plot for use with SAQ 4.5.

Objectives for Chapter 4

After studying Chapter 4 (and any associated audio, video or TV material), you should be able to:

- 4.1 Give brief definitions of the terms, concepts and principles listed below.
- 4.2 Summarize the main criteria used to identify craters as being of impact origin.
- 4.3 Describe the morphology and structure of a pristine crater, and the effects of variations in scale, surface gravity, angle of incidence and target materials.
- 4.4 Given scaled spacecraft images of unknown planetary surfaces, make inferences about the sizes of craters present and the properties of the planet or satellite.
- 4.5 Summarize the range of variables influencing the size-frequency distribution of craters on a planetary surface.
- 4.6 Given suitable data, plot an appropriate crater size frequency distribution graph for a planetary surface, and use it to draw inferences on the cratering history of the object.
- 4.7 By examining plots for different planetary bodies, draw correct inferences about their different relative ages and cratering histories, and use lunar data to estimate absolute ages for specimen surfaces.
- 4.8 Explain how 'adjusted' lunar data can be used to estimate the absolute ages of unknown planetary surfaces.
- 4.9 Given photographs or data on unknown craters or crater fields, comment on their mechanism of origin and size-frequency distribution, and how these relate to the nature and history of the planetary body concerned.

List of scientific terms, concepts and principles used in Chapter 4

Term	Page	Term	Page	Term	Page
central peaks	80	impact melt	73	rays	73
continuous ejecta	73	mare (maria)	80	saturation	84
crater production flux	84	Meteor Crater	71	secondary craters	73
discontinuous ejecta	73	multi-ringed basin	80	shock wave	72
ejecta	72	normal distribution	83	size-frequency distribution	83
glass	73	palimpsest	78	tektites	73
hyper-velocity impacts	72	peak-ring basin	80	terraces	80
impact cratering	69	rampart craters	76	transient cavity	72
impact flux	83	rarefaction wave	72	viscous relaxation	78

Chapter 5

Volcanism

Prepared for the Course Team by Peter Francis

5.1	Introduction	95
5.2	Melting and magma production	95
5.2.1	Partial melting of chondritic silicates	96
5.2.2	Partial melting in icy satellites	97
5.2.3	Thermal properties of ices and implications for volcanism	97
	Summary of Section 5.2 and SAQ	98
5.3	Styles of volcanism	99
5.3.1	Effects of gravity	99
5.3.2	Effusive and explosive eruptions	99
5.4	Lavas	99
5.4.1	Viscosity and lava flow thicknesses	101
5.5	Pyroclastic eruptions	102
5.5.1	Dynamics and dimensions of eruption columns	103
5.5.2	Height of eruption columns	103
5.5.3	Dispersal of volcanic ejecta	104
5.5.4	Formation of pyroclastic flows	105
	Summary of Sections 5.3–5.5 and SAQs	107
5.6	Volcanism in the Solar System	108
5.6.1	Earth	108
5.6.2	The Moon	109
5.6.3	Mercury	110
5.6.4	Venus	110
5.6.5	Mars	112
5.6.6	Io	115
	Summary of Section 5.6 and SAQs	118
	Objectives for Chapter 5	118

5.1 Introduction

In the two decades after the Apollo landings on the Moon, unmanned spacecraft returned a flood of data from Mercury, Venus, Mars, Jupiter, Saturn, Uranus and Neptune. For volcanologists, one of the most exhilarating aspects of these historic missions was the discovery that volcanism is a Solar System-wide process. They can no longer sit back cosily contemplating a single planet, but must now think about three other planets, and a squadron of satellites. Previously, they had been confronted only by the awesome but familiar aspects of 'hot' volcanism, driven by melted rocks. Now they must also deal with the chilly processes of icy volcanism, driven by water and other more exotic compounds.

Fortunately, the laws of physics are indeed universal. In this chapter, therefore, we aim to provide you with sufficient understanding of the underlying causes of volcanism and of the physics of volcanic eruptions for you to be in a position to predict the nature of the volcanic phenomena that you would expect to find on even the most obscure planetary body, given only basic information about its size, mass, density and orbital parameters.

There are two basic requirements for any sort of volcanism: a source of heat, and something to melt. You encountered the various sources of heat and the mechanisms of heat transfer in planetary bodies in Section 3.5. Before reading on, check your understanding of these important topics by attempting the SAQs which follow. If you have difficulty with them, go back and re-read the appropriate parts of Section 3.5.

SAQ 5.1 (Objective 5.5) Would you expect there to be any heat generated by tidal dissipation in the Moon?

SAQ 5.2 (Objective 5.5) A hypothetical planet with the same mass as the Earth is made completely of pure water-ice. What sources of internal heat might it have?

SAQ 5.3 (Objective 5.5) Consider an Earth-like planet which is losing heat from its interior today by conduction through its lithosphere, hot-spot volcanism (advection) and plate recycling in equal proportions. How would you expect its heat loss mechanism to evolve in the future? Sketch a possible evolutionary path on a triangular diagram like that shown in Figure 3.14, and write a sentence or two to justify your conclusions.

5.2 Melting and magma production

We now turn to the second fundamental ingredient of volcanism, the melting of solid materials to produce molten material, or magma. This requires an examination of the raw materials involved, the compositions of planets and satellites.

Meteorites have long been used as pointers to the composition of the planets. They will be examined more fully in Chapter 8, but for now, it is sufficient to note that early ideas for the bulk composition of the Earth simply assumed that it was the same as that of a group of silicate meteorites known as CI carbonaceous chondrites, a view enshrined as the **chondritic Earth model**.

Various models have been proposed for the compositions of the rocky, terrestrial planets. At present, the best one can say is that the terrestrial planets are of approximately chondritic composition, but some are more chondritic than others. Mars, for example, may have a ratio of potassium to uranium which is more nearly like that in chondrites than the Earth.

Recall from Subsection 2.2.3 that 'chondritic composition' means chemical elements are in the same proportions as in carbonaceous chondrites.

□ **Why should this particular ratio be important?**

■ Both potassium and uranium have radioactive isotopes with different half lives (Table 3.2), so the ratio of these elements influences the planet's potential **heat production and thermal evolution**.

Most of the satellites of the giant planets are made predominantly of water ice, mixed with increasingly more volatile compounds such as ammonia and methane at Uranus and beyond (Section 2.6). Reference to Figure 1.4 shows that in terms of density (and size) the solid bodies of the Solar System fall into ill-defined groups: there are many objects which have densities intermediate between icy and rocky. Thus, planetary volcanism involves melting of two distinctly different raw materials: high temperature melting of silicates, and low temperature melting of ices, sometimes termed **cryovolcanism** (from the Greek *kryos*, meaning frost) to distinguish it from the familiar hot variety.

5.2.1 Partial melting of chondritic silicates

Much the commonest products of the melting of silicates in a body of chondritic composition are **basalt** lavas – dark, fine grained rocks which contain about 50% SiO_2 . Basalt lavas are almost ubiquitous on rocky bodies – they cover 70% of the Earth's surface, and comparably huge areas on Mars and Venus. It is not difficult to see why basalts are so abundant: all the terrestrial planets have broadly similar chondritic compositions, and basalt is what results from melting chondritic silicates. Of course, if material with the composition of a chondritic meteorite were *completely* melted, then the melt would have the same composition as the meteorite. But this is not what happens in nature. When mixtures of minerals are heated, some components melt at lower temperatures than others, in a process known as **partial melting**.

Partial melting is a vitally important process in generating magmas. When partial melting takes place, elements that are 'incompatible' with the source rock composition are mobilized and escape into the melt. Typically, **incompatible elements** have large ionic radii – potassium is an outstanding example. Other, 'compatible' elements, such as iron and magnesium tend to remain in the solid **residue in minerals stable at high temperatures, which include olivine and pyroxene**. The smaller the extent of partial melting, the more fractionation (separation) takes place between compatible and incompatible elements. A kind of concentration process therefore takes place during partial melting, giving rise to liquids richer in SiO_2 – potassium and other incompatible elements than the source rocks. Similar fractionation took place on a planetary scale during the **formation of the Solar System: some elements were concentrated into the metallic cores of planetary bodies, and others were concentrated into their crusts** (Subsection 3.3.2).

On Earth, the basalts which cover 70% of its surface form the oceanic crust and are termed **mid-ocean ridge basalts** or MORBs. They form in such huge **volumes at ocean ridges that it was supposed for a long while that they result from direct melting of mantle peridotite, and that they are therefore 'primitive', i.e. derived directly from nearly chondritic materials.** This reasonable supposition is now known to be wrong: MORBs form by multi-stage processes and are not primitive at all. At present, the most primitive known basalts are some scarce lavas, **richer in magnesium than MORBs, which contain nodules of mantle materials.** Some primitive basalt lavas are also found in quite different contexts: in ancient rocks, more than 2 billion years old. These lavas, known as **komatiites** after a locality in South Africa, provide important pointers to the thermal state of the Earth in its early history. Basalts with compositions broadly similar to komatiites are widespread on the other silicate planets (Table 5.1).

Table 5.1 Compositions of fresh (unaltered) mantle and some important basalt types (% by mass)

Component	Fresh mantle	Primitive basalt	MORB	Ancient terrestrial lava (komatiite)	Lunar mare basalt	Shergotty meteorite (Mars basalt)
SiO ₂	45.1	44.2	50.5	45.2	43.6	50.1
TiO ₂	0.2	3.7	1.6	0.2	2.6	0.9
Al ₂ O ₃	3.3	12.1	15.3	3.7	7.9	6.7
FeO	8.0	10.9	10.5	11.0	21.7	18.7
MnO	0.1	0.2	—	0.2	0.3	0.5
MgO	38.1	13.1	7.5	32.2	14.9	9.4
CaO	3.1	10.1	11.5	5.3	8.3	10.0
Na ₂ O	0.4	3.6	2.6	0.4	0.2	1.3
K ₂ O	0.03	1.3	0.2	0.1	0.1	0.2
P ₂ O ₅	—	0.8	0.1	0.02	—	0.7
Cr ₂ O ₃	0.4	0.1	—	—	0.9	0.2

11Q 5.1 Use the data in Table 5.1 to compare the compositions of the komatiite lava, lunar mare and Martian lavas. Which has the most primitive composition and which is most MORB-like? Account for the differences between them in a sentence or two.

5.2.2 Partial melting in icy satellites

If a satellite were wholly composed of water-ice, then any ‘magma’ that resulted from melting would be pure water. However, most icy satellites contain many other components, including frozen gases such as methane and ammonia, and dissolved salts derived from silicate rocks, such as magnesium sulphate.

What effect might such dissolved salts have on the melting temperature of the ice concerned?

■ As happens when ordinary salt is mixed with water, the melting temperature would be reduced several degrees, making melting easier.

Mixing water with more volatile compounds such as ammonia or methane produces more important effects. An ice made of ammonia and water begins to melt at temperatures about 100 K lower than that for pure water-ice. Partial melting of such an ice would yield a liquid much richer in ammonia than the solid mixture. However, most icy satellites are so small and contain so little rock in which heat-producing radioactive elements could occur, that their internal temperatures would be so cold – of the order of 100 K – that melting near the surface at the present day would be inconceivable were it not for heat generated by tidal dissipation. There is evidence that tidal melt generation has happened in the past on some icy satellites of Saturn and Uranus (Plates 2.9–2.11, 2.21).

5.2.3 Thermal properties of ices and implications for volcanism

The **specific heat capacity** (the amount of heat required to raise the temperature of 1 kg by 1 K) of basaltic rocks is about $1.2 \times 10^3 \text{ J kg}^{-1} \text{ K}^{-1}$, and their **latent heat of fusion** (the amount of heat required to melt 1 kg completely *after* it has been brought to its melting temperature) about $4.8 \times 10^5 \text{ J kg}^{-1}$. Thus, to raise

one kilogram of basaltic rock from room temperature (say 290 K) to its melting temperature (say 1 500 K) requires $1.2 \times 10^3 \times (1\,500 - 290)$ J and then to melt it completely requires about a further 4.8×10^5 J. For comparison, the specific heat capacity of water-ice is $2.1 \times 10^3 \text{ J kg}^{-1} \text{ K}^{-1}$ and its latent heat of fusion $3.34 \times 10^5 \text{ J kg}^{-1}$.

The latent heat of fusion of ice is similar to that of rock. Does this mean that a given input of thermal energy, for example from tidal dissipation, should produce similar amounts of melting within small rocky and icy satellites?

■ No - the amount of heat required to bring the rock or ice to its melting temperature also has to be taken into consideration, as well as that required for melting. Because silicates melt at far higher temperatures ($\approx 1\,500$ K) than water-ice (273 K), far more heat is required to cause melting in a rocky satellite than in an icy satellite of similar mass.

Whether cryovolcanism can occur on an icy satellite depends on a balance between the extremely low surface temperatures in the outer Solar System, and the amount of heat generated by tidal interactions – augmented by a significant amount of radiogenic heat (Subsection 3.3.2) only for large satellites containing a substantial proportion of rock. Many icy satellites show no evidence of cryovolcanism at all, but several others do. Europa (Plate 2.22), the galilean satellite next out from Io, experiences large tidal forces in its resonant orbit, and its icy surface is young and almost unblemished by craters. Icy volcanism has clearly taken place in the recent past. However, Europa is dominantly a silicate body, with only an outer skin of ice, and its ‘background’ temperature is probably controlled by radiogenic heating.

Enceladus, a satellite of Saturn only 251 km in radius shows some of the best evidence of volcanism on a dominantly icy body. Its surface shows evidence of a long and complex geological history (Plate 2.21). Some regions are heavily cratered, others have been resurfaced probably by cryovolcanism. Unfortunately, our knowledge of cratering statistics in the vicinity of Saturn is not good enough to assign ages confidently to these two types of surface, but one is obviously relatively old; the other relatively young.

11Q 5.2 Because Enceladus is so small, and mainly icy, its volcanic processes must be almost exclusively tidally driven. Inspect the data on planetary satellites in Table 1.1b. Is there any evidence in these data that lends support to this hypothesis?

Summary of Section 5.2 and SAQs

- 1 Volcanism requires two ingredients: a source of heat, and something to melt.
- 2 This section reminded you of the ways outlined in Chapter 3 in which heat is generated and transported to the surface of planetary bodies.
- 3 There are two main kinds of raw material for melting in the Solar System: silicates, with compositions approximately represented by carbonaceous chondrite meteorites; and ices, mostly water, but with a greater proportion of even more volatile substances at greater distances from the Sun. Basalt lavas of varying compositions are the Solar System-wide result of partial melting of chondritic silicates.
- 4 Although ices require less thermal energy for melting than silicates, most icy satellites are too small to have experienced much melting, unless this is driven by dissipation of tidal energy.

SAQ 5.4 (Objective 5.2) Refer to Table 5.1, which summarizes the compositions of basalt lavas on various planetary bodies. Which of these lavas do you think is likely to approximate most closely to the composition of a basalt on Venus, and why?

5.3 Styles of volcanism

The way in which a magma makes its way to the surface of a planetary body, and the nature of the resulting eruption are controlled by both intrinsic and extrinsic physical factors. Intrinsic factors are those governed by the stuff that melts and the magma which is formed: its temperature, composition, and content of dissolved gases. Extrinsic factors are dictated by the planetary environment, especially gravity and atmospheric pressure.

5.3.1 Effects of gravity

Gravity provides the buoyant force that causes a magma to rise through the crust of a planet. Consider a magma that is slightly less dense than the crust surrounding it. The magnitude of the buoyant force acting on the magma depends on how big the magma body is, how much less dense it is than the surrounding material, and gravity.

Will the buoyant force be greatest in a planetary body with high or low gravity?

■ The force will be greatest in a high-gravity body [*Comment.* Think about a cork floating on water. If gravity were reduced to zero, by doing the experiment on the orbiting Space Shuttle, the buoyant force on the cork would be zero, and both cork and water would drift weightless around the cabin.]

While gravity is necessary, it is not the only factor. To reach the surface, a buoyant blob of magma has to overcome forces resisting its ascent through the lithosphere. These resisting forces are not gravity-controlled, but are mostly dictated by the intrinsic properties of the magma, such as its viscosity. Squeezing toothpaste out of a tube requires the same effort everywhere, whether in the bathroom at home or aboard the Space Shuttle.

5.3.2 Effusive and explosive eruptions

Volcanologists recognize two contrasted styles of volcanism: *effusive* and *explosive*. Effusive eruptions yield lava flows; explosive ones usually generate fragmentary **pyroclastic rocks** (Greek, 'fire-broken'), colloquially known as volcanic *ash*. Effusion of lava may be accompanied by minor pyroclastic activity, but it is often so tranquil that observers can stand within a metre or two of vents from which molten silicate rock is pouring at several cubic metres per second. **Pyroclastic eruptions**, by contrast, can be exceptionally violent, and have been associated with some of the greatest natural catastrophes in Earth history. Explosions from the 1883 eruption of Krakatau (Indonesia) were audible over much of Australia and the Indian Ocean, further than the reverberations of even the largest nuclear weapons test explosions.

A pyroclastic eruption is one that produces pyroclastic rocks.

Krakatau, pronounced 'crack-at-ow', used to be known incorrectly as Krakatoa.

5.4 Lavas

Differences in magma composition and gas content are vitally important determinants of the variations between extremes of quiet lava effusion and catastrophic explosion. Magma **viscosity** is also vitally important. Viscosity is a

property of fluids that is familiar from everyday life, but is more difficult to discuss quantitatively. For our purposes, it can be defined as *the internal resistance to flow by a substance when a stress is applied*. In the simplest case, this stress may merely be that which acts on a liquid when the bottle containing it is tilted.

Lava viscosity is heavily dependent on composition: other things being equal, the more SiO_2 a lava has, the more viscous it will be. From a planetary perspective, this compositional variation is relatively unimportant, because lavas are mainly of basaltic composition. Earth seems to be a special case, since lavas with a huge range of compositions, and thus viscosities, are erupted.

In lavas, as in other fluids like motor oil, viscosity decreases rapidly with increasing temperature. Basalts are usually erupted at temperatures of about $1\,200\text{ }^\circ\text{C}$. The only lavas which are hotter and of lower viscosity are primitive iron- and magnesium-rich basalts (like komatiites, Table 5.1) which may be erupted at temperatures of more than $1\,400\text{ }^\circ\text{C}$.

Basalts of identical compositions erupted at identical temperatures may nevertheless give rise to land-forms ranging from flat sheets to tall conical volcanoes. Can you suggest another important parameter, other than viscosity, that influences the shape of lava accumulations?

■ It is **eruption rate**. When basalt lavas ooze slowly to the surface, they chill quickly, and small flows pile up on top of one another. At high eruption rates, such large masses of lavas are erupted so quickly that their heat content is sufficient to enable them to flow large distances before chilling.

This explains an apparent paradox of lunar volcanism: lavas are abundant, but there are no obvious volcanoes. After massive impacts formed the multi-ringed mare basins (Subsection 4.3.2), basalt lavas welled up in huge volumes from the lunar mantle, flooding the basins like bath tubs filled from beneath (Figure 5.1). Similar torrents of basalt lavas flooded parts of Mercury, Venus and Mars. On Earth, there is nothing exactly comparable to the lunar maria, but there are many areas of **continental flood basalts**, where huge volumes of basalt spewed to the surface to form extensive flat plateaux. One example is the 65 million year old Deccan province of northern India and Pakistan (Figure 5.2a). Some scientists believe that eruption of the Deccan lavas 65 million years ago was implicated in the death of the dinosaurs. Another example is the 15 million year old Columbia River basalt province of North America (Figure 5.2b): over 700 km^3 of basalt gushed to the surface in the Columbia River Plateau to form the Roza Flow, which travelled more than 300 km westwards along the Columbia River Gorge. One estimate suggests that the lava flow front was 30 m high, extended laterally for over 100 km, flowed forward at 5 km h^{-1} , and the eruption rate so great that the whole volume may have been erupted in a week. Recent work suggests that colossal outpourings of flood basalts 245 million years ago in Siberia may have been linked to the greatest mass extinction event in Earth history, which defines the boundary between the Paleozoic ('early life') and Mesozoic ('middle life') Eras.

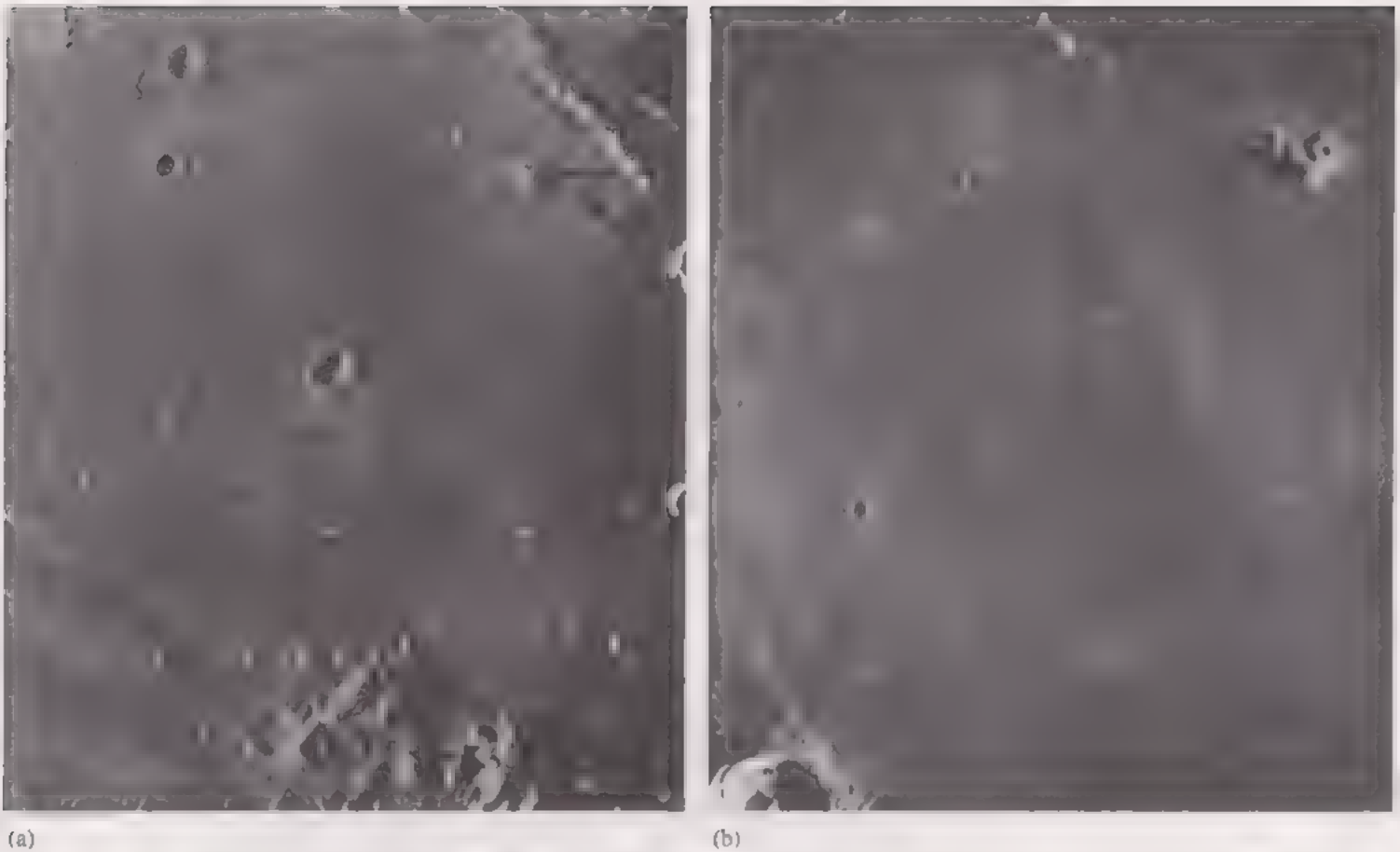


Figure 5.1 (a) A view of the smooth basalt plains filling the Mare Imbrium, obtained with the 64 cm telescope at the Pic du Midi in France. This is as good an image as is possible to obtain from the Earth, but it is impossible to distinguish any volcanic features. The crater just above the centre is Euler, 25 km in diameter. (b) A detailed view of part of the same area, obtained from an Apollo spacecraft. Euler (an impact crater) is at the bottom left. At this resolution, the secondary craters around Euler are easily seen, and two lava flows extend towards the top of the picture, passing between the isolated hills. There is no obvious volcano as a source of the flows.

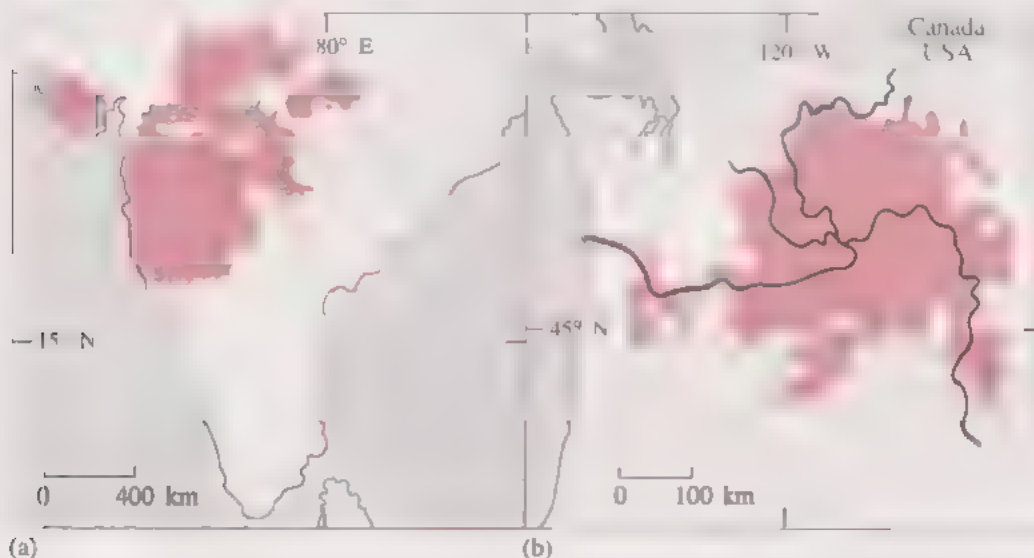


Figure 5.2 Sketch maps showing the distributions of (a) the Deccan and (b) Columbia River basalt provinces. Note the difference in scale between the two sketches. In both cases, lava accumulated to a total of 1–2 km in thickness

5.4.1 Viscosity and lava flow thicknesses

A blob of viscous liquid such as honey may seem superficially to behave like lava when it oozes over a flat surface. Over time, however, a blob of honey flattens out to a thin film: it has little strength to resist spreading. Two forces

You are not expected to remember Equation 5.1 (or the subsequent Equations 5.2 and 5.3); however, you should note that each has a reasonable general form. For example, in Equation 5.1, the greater the value of g (the stronger the gravity), the smaller the value of h , as you might expect.

influence the rate at which a blob spreads: gravity, and viscous resistance. The magnitude of the gravitational force depends, of course, on the mass and radius of the planet that our blob of honey is oozing stickily over, but viscosity is a property intrinsic to the material. In addition to viscous resistance, molten lavas also possess a **yield strength** partly derived from the chilled crust which forms on their surface. Before a blob of lava can spread, its yield strength must be overcome by the forces driving flow. Thus, lavas never form thin films, they always retain some thickness, which is a measure of their yield strength. Fluid basaltic lavas such as those of Hawaii commonly flow in oozy dribbles as little as a few centimetres thick; by contrast rhyolites (containing about 70% SiO_2) form flows which are never less than many metres thick. The critical thickness, h , that a body of lava resting on a sloping surface must obtain before it will begin to flow is given by a simple equation:

$$h = \frac{S_y}{g \rho \tan \alpha} \quad (5.1)$$

where g is the acceleration due to gravity (at the surface); ρ is the lava density; S_y is the lava yield strength; and α is the angle of slope from the horizontal.

Once the flow gets moving, things get complicated, because of the large number of variables. Apart from temperature and composition, lava viscosity is also a function of the *rate* of movement. Thus, a basalt flow moving on a steep slope, where internal deformation is rapid, will have a lower viscosity than one moving on a gentle slope, other factors being equal.

11Q 5.3 (a) Basaltic lavas of similar composition are found on Mercury, the Moon, Mars and some asteroids. Where would you expect to find the thickest lava flows, and why? (b) A basaltic magma gives rise to a lava flow 10 m thick on the Moon. How thick would the flow be on Mercury, other circumstances such as eruption rate being similar? (Refer to Table 1.1a.)

5.5 Pyroclastic eruptions

Gases in a magma are often referred to as 'volatiles'. Common volatiles in silicate magmas include water, sulphur dioxide and carbon dioxide. An example of a volatile present in water-dominated cryovolcanic magmas is methane. You will again meet gases described as volatile in the context of planetary atmospheres in Chapter 6. Note that this is a rather different usage of the term 'volatile' from that in Chapters 2 and 3, where the processes that concerned us involved condensation in the solar nebula.

When a magma rises to the surface of a planet, it contains dissolved gases; usually water vapour, carbon dioxide, and sulphur dioxide. Basaltic magmas rarely contain more than 1% by mass of dissolved gases, but none the less, these escaping gases can spray jets of incandescent magma hundreds of metres high. Such **fire fountains** are typical facets of eruptions in the Hawaiian islands, and are amongst the Earth's most spectacular natural phenomena (Plate 2.25).

Given that they form such a small component by mass, can you suggest two reasons to explain how escaping gases can play such important roles in basaltic eruptions?

■ Dissolved gases, being low in density, may accumulate at the top of the magma reservoir. More importantly, because gases have molecular masses which are much lower than those of the materials in the magma, there are actually large *numbers* of gas molecules, even though these form only a small fraction of the total *mass*.

Explosive eruptions are driven by the thermal energy stored in the magma. Thermal energy is transferred into the kinetic energy of the ejected material through the expansion of the gases diffusing into growing bubbles. Magma viscosity and dissolved gas content are therefore the pre-eminent variables influencing the course of an explosive eruption. Because the laws of physics really are universal, it is possible to predict the patterns of pyroclastic eruptions on planetary bodies other than the Earth. Under low gravity, bubbles will begin

to form at greater depths than on the Earth, encouraging efficient volatile release. Magmas containing even small quantities of dissolved volatiles will be sprayed from volcanoes in fire fountains.

However, pyroclastic activity can only take place at all if there is sufficient dissolved gas to expand and disrupt an ascending magma. On Earth, water vapour is the most important gas driving volcanic eruptions (Table 5.2). The amount of gas expansion is greatest when atmospheric pressure is least. As a result, on Mars, where atmospheric pressure is only 0.7% that of the Earth, a gas content of only $\approx 0.01\%$ by mass is required to drive pyroclastic eruptions; for Earth the minimum is $\approx 0.07\%$; while for Venus, no less than 3% is required to overcome the crushing atmospheric pressure (90 times the Earth's). On Mars, there is abundant evidence of volatiles (such as the ice in its polar caps), so pyroclastic eruptions should be favoured. On hot, dry Venus things are more problematic. Bearing in mind that terrestrial basalts typically contain less than 1% volatiles (mostly water), it is difficult to conceive that gas concentrations of as much as 3% could be achieved in Venus's magmas. Possibly, carbon dioxide may play a role.

5.5.1 Dynamics and dimensions of eruption columns

On bodies without atmospheres, pyroclastic eruptions are straightforward – material is simply sprayed into space, and falls back along parabolic paths. When small eruptions take place in an atmosphere, the situation remains fairly simple, the particles falling back along paths modified by atmospheric drag. These paths can be calculated from ballistic theory, refined by generations of artillery officers in the intervals between shelling each other.

Large eruptions are much more complex. They may propel towering **eruption columns** tens of kilometres into the atmosphere. Like Caesar's Gaul, large eruption columns can be divided into three parts. Immediately above the vent is a **gas thrust region**, in which particles are propelled upwards like shot from a gun. Above the gas thrust region is a **convective ascent region**, in which hot, buoyant gases carry smaller fragments upwards, while larger fragments fall back. Convection of hot magmatic gases is aided by heat transfer between small, hot magma fragments and entrained air – large amounts of cool atmospheric gases are drawn into the column, to be heated by the magmatic material within it. At high altitude, the plume becomes neutrally buoyant (it neither sinks nor rises), and spreads out to form an **umbrella region** (Figure 5.3).

5.5.2 Height of eruption columns

Hot air balloons drifting like gaudy bubbles on still summer evenings demonstrate that hot gas is buoyant. A major eruption column may rise tens of kilometres and may be sustained for many hours. It will rise convectively until it reaches a level where its density is the same as the surroundings, so it is no longer buoyant. At this level, H_b , it spreads sideways (Figure 5.3). Some material in the centre rises higher, to H_T in Figure 5.3, carried further upwards by momentum, not buoyancy. This gives rises to a plume resembling the mushroom cloud of nuclear explosions (although these explosions are the results of single thermal pulses, rather than sustained convection). Decades of studies of the thermodynamics of hot gases have shown that the height H_T which plumes will reach can be estimated from the relationship:

$$H_T = k(M\Delta T)^{1/4} \quad (5.2)$$

where M is the mass eruption rate of magma; ΔT is the difference in temperature between the newly erupted magma and the surrounding atmosphere at ground level; and k is a constant. The mass eruption rate is critical because it is a

The Earth's 'Standard' atmospheric pressure is 1 013.25 millibars (mbar) at sea-level. 1 bar = 10^5 Pa, so 1 mbar = 10^2 Pa

Table 5.2 Proportions of principal gases in a typical basaltic volcano (Kilauea, Hawaii)

Gas	Volume (%)
H ₂ O	56.2
CO ₂	27.9
SO ₂	13.6
H ₂ S	0.2
HCl	0.2
H ₂	0.9
CO	1.0

Among other things, k includes atmospheric properties.

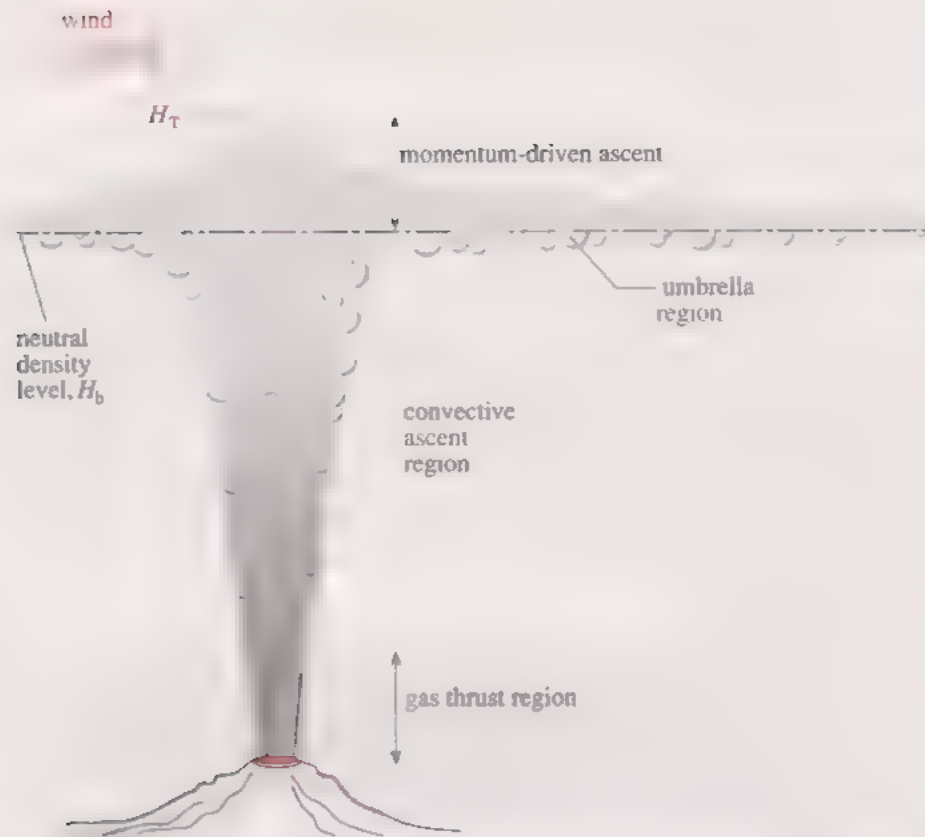


Figure 5.3 Components of a volcanic eruption column rising through an atmosphere. In large eruptions on Earth, the convective ascent region may reach more than 40 km vertically, while the umbrella region may extend hundreds of kilometres downwind. A small amount of upwind spreading also takes place.

measure of the *thermal energy* being pumped into the eruption column. Essentially, the height reached varies as the fourth root of the eruption rate, other factors being equal.

What change in eruption rate would be required to *double* the height reached by a convecting column?

■ A 16-fold increase would be needed.

5.5.3 Dispersal of volcanic ejecta

When a large eruption column propels solid fragments high into an atmosphere, the distribution of the volcanic ejecta over the surface is controlled by the wind, and by the **terminal fall speed** of the particles in the eruption cloud. Terminal fall speed is a simple concept, one that can readily be measured in air by the simple means of dropping objects from a great height and timing their fall. A free-fall parachutist (a large, soft object) reaches a terminal speed of about 60 m s^{-1} . Terminal fall speeds can be estimated from the equation:

$$\text{terminal speed} = C_d \sqrt{\frac{dg\rho}{\beta}} \quad (5.3)$$

where C_d is the **aerodynamic drag coefficient**, d is fragment diameter, g is the acceleration due to gravity; ρ is the fragment density; and β the atmospheric density.

Motor cars and aircraft have a parameter similar to C_d . Volcanic ash fragments are, of course, far removed from the smoothness of luxury cars – they are irregular in size and shape and have rough, drag inducing surfaces. Values of C_d of about 1.054 have been found in experiments.

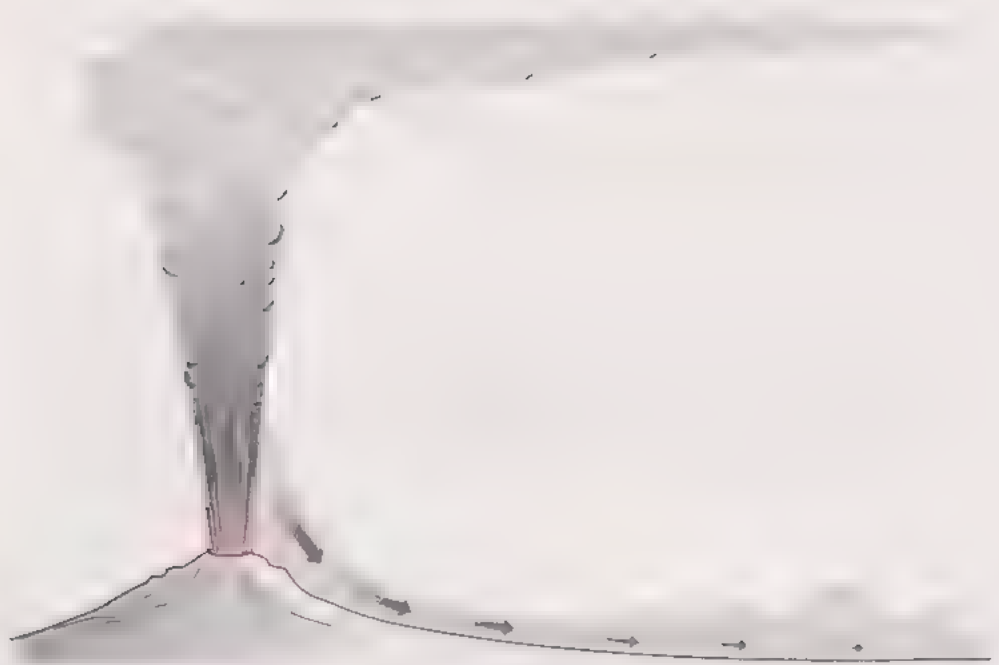


Figure 5.5 Pyroclastic flows form by collapse from an eruption column which is so dense that convective rise is inhibited. Descending material acquires enormous kinetic energy and travels over the ground as avalanches of incandescent pumice and dust. Exceptionally, flows on Earth may travel over 100 km.

Such large eruptions have fortunately been rare in recent Earth history. Eruption of more than $2\,000\text{ km}^3$ of pumice 75 000 years ago formed the 100 km long Toba caldera in Indonesia, now filled with a lake (Figure 5.6). Although most of the magma was erupted in the form of pyroclastic flows, ash fell as far distant as north India.

Two independent parameters are important in controlling the mass eruption rate, and hence the density and stability of an eruption column. They are the volatile (gas) content of the magma and the vent radius.

Magmatic volatiles drive all pyroclastic eruptions; without volatiles a magma would be as flat as warm British beer. More volatiles means more fizz. Other things being equal, this yields higher eruption speeds and therefore higher mass eruption rates. For a given vent radius, model calculations indicate that eruption speed increases sharply as volatile content increases. Magmas on Earth with SiO_2 contents between 65 and 70% typically have 2–3% volatiles. According to the model in Figure 5.7a, they should achieve speeds of up to 400 m s^{-1} when erupted through vents 330 m in diameter. If the volatile content in

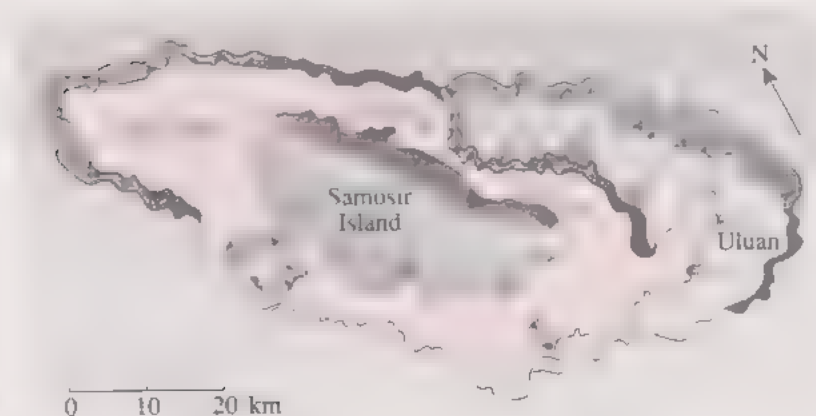


Figure 5.6 Topographic sketch of the giant Toba caldera, Sumatra, which is 100 km long. The caldera consists of a vast, elongate depression, now filled by Lake Toba, with a central elevation (Samosir Island) marking a site where later magmatic movements built the caldera floor up by about 500 m.

the model is held fixed (Figure 5.7b), and the vent radius increased, the result is an increase in eruption speed. This is intuitively reasonable: it is easier to blast gas and magma up through a wide pipe than a narrow one. Eruption speed translates directly into mass eruption rate, and into thermal energy, which is the chief control on eruption column height, as in Equation 5.2. Thus, increasing magmatic volatile content and vent radius both cause increasing column height.

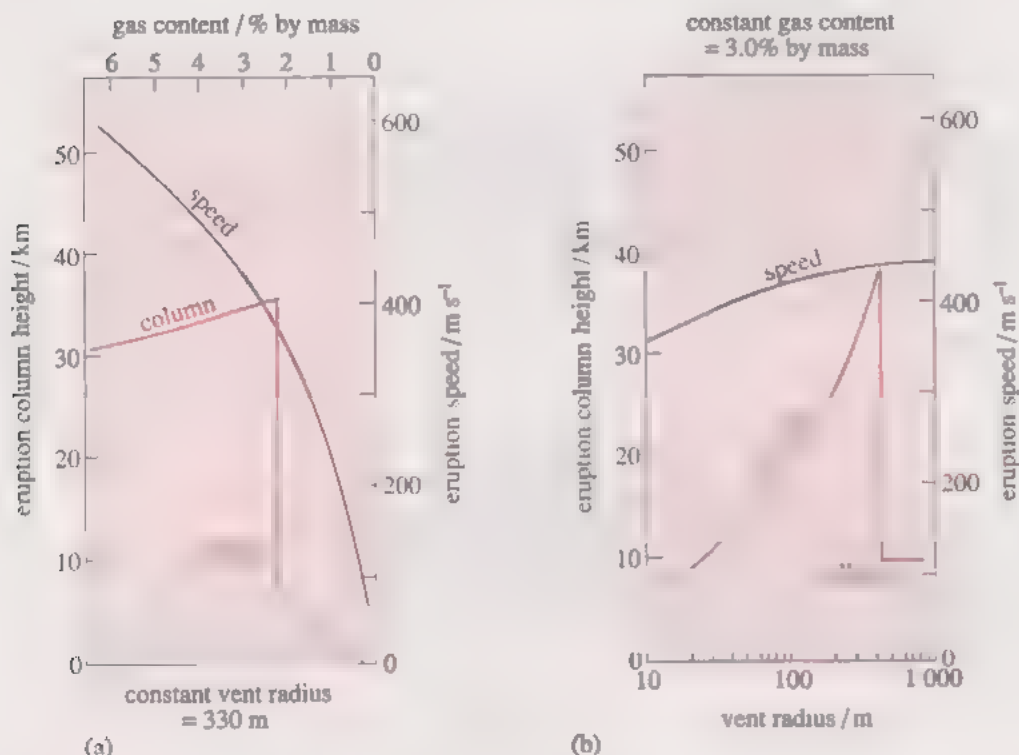


Figure 5.7 (a) Graph showing the relationship between eruption speed and volatile content for a volcanic vent of fixed radius (330 m; black line). The red line shows the height of the resulting eruption column. (b) Graph showing relationship between eruption speed and vent radius for a volcanic vent with fixed volatile content (3.0% by mass). The solid line again shows the height reached by the eruption column.

11Q 5.5 Would you expect to find convecting eruption columns on small bodies such as the Moon that lack atmospheres? If not, how would you expect the products of pyroclastic eruptions to be dispersed?

Summary of Section 2.2 and SAQs

- 1 Volcanism may be either effusive or explosive. Effusive volcanism yields lava flows, which may accumulate to form mountains (volcanoes) or flood wide areas to form lava plains and plateaux. Explosive volcanism is driven by the explosive escape of volatiles, and yields pyroclastic fall and flow deposits, whose form is influenced by atmospheric pressure.
- 2 In the absence of an atmosphere, ejected particles follow ballistic trajectories, while an atmosphere ensures widespread dispersal of fine-grained material by winds.
- 3 Pyroclastic flows extend radially around the volcanic vent, reaching hundreds of kilometres.
- 4 The sites of large explosive eruptions are generally depressions (calderas) surrounded by extensive aprons of pyroclastic fall and flow deposits.

SAQ 5.5 (Objectives 5.3 and 5.4) A lump of pumice erupted by a convecting eruption column on Earth has a terminal fall speed of 100 m s^{-1} . How fast would it fall on Venus, to a first approximation?

SAQ 5.6 (Objectives 5.3 and 5.4) Magma with a volatile content of about 1.25% is erupted through a vent on the Moon which is 200 m in radius at a speed of 250 m s^{-1} , and a mass eruption rate of about $3 \times 10^8 \text{ kg s}^{-1}$. What sort of eruption would you expect?

SAQ 5.7 (Objectives 5.3 and 5.4) Consider pyroclastic eruption columns on the Earth and Mars that carry material up to an altitude of 25 km. If a steady wind of 20 m s^{-1} were blowing in both cases, how would the shape of the resulting isopleths on Mars compare with those on Earth, and why?

5.6 *Volcanism in the Solar System*

So far in this chapter, we have examined the physical processes of volcanism. To conclude, we examine briefly the evidence for volcanism in the Solar System, and the contrasts between the volcanic histories of different kinds of bodies. We focus on the observable evidence for different volcanic phenomena.

5.6.1 *Earth*

Terrestrial volcanism is an on-going process. 70% of the Earth's surface is covered by basalt lavas erupted within the last 200 million years, although most of them are hidden from sight beneath the oceans. Several hundred active volcanoes around the world provide conclusive evidence of continuing thermal processes within the mantle. Both effusive and explosive eruptions are common. Although they are not the world's highest mountains, Mauna Loa and Mauna Kea in Hawaii are the largest, reaching 9 km above their bases on the ocean floor.

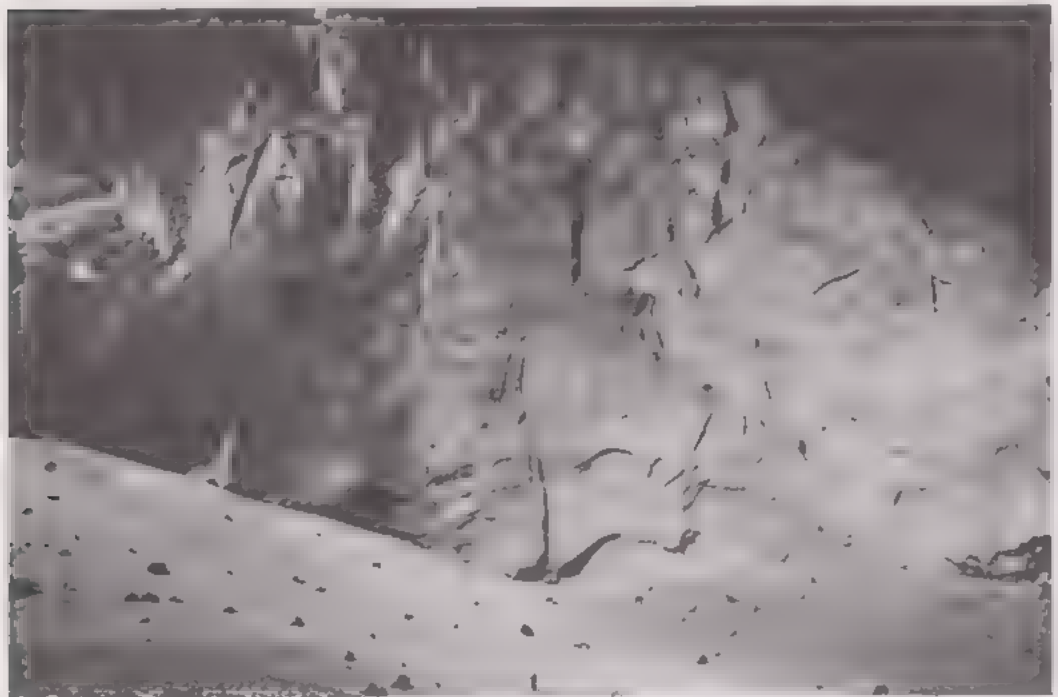


Figure 5.8 The geologist at the bottom of this photograph, taken in the western flanks of Cerro Galan (in northwest Argentina), illustrates the enormous scale of terrestrial pyroclastic eruptions. Most of the pyroclastic flows making up these 300 m high cliffs here were erupted during a single event, which may have lasted only a few hours.

In the Deccan area of northwest India, basalt flows erupted 65 million years ago total more than 1 km in thickness and cover almost a million km². In parts of northern Mexico and the Central Andes, pyroclastic flow deposits cover hundreds of thousands of km², and have shaped the landscape (Plate 2.24; Figure 5.8).

5.6.2 The Moon

As discussed in Sections 4.3 and 4.4, basalt lava flows form the youngest extensive rocks on the Moon, some of them 3 200 million years old, filling the mare basins. Spacecraft photographs (Figure 5.1) reveal the outlines of many flows, some of them extending hundreds of kilometres, but the sources of these flows remain obscure. There are *no* obvious volcanoes – the lavas must have flooded to the surface via relatively narrow cracks and fissures.

Because the Moon is extremely deficient in volatiles (Table 3.1), including elements that would form volcanic gases, large-scale pyroclastic activity is not to be expected. None the less, evidence for some pyroclastic activity has been detected on spacecraft images, and in surface samples.

Given the absence of an atmosphere on the Moon and its low surface gravity, how would you expect pyroclastic materials to be dispersed, and what might the deposits look like?

■ Explosive eruptions taking place into the vacuum of space on a low-gravity planetary body would propel ejecta long distances along parabolic (ballistic) trajectories. The deposited material would form thin sheets around the vent, with larger fragments concentrated towards the vent.

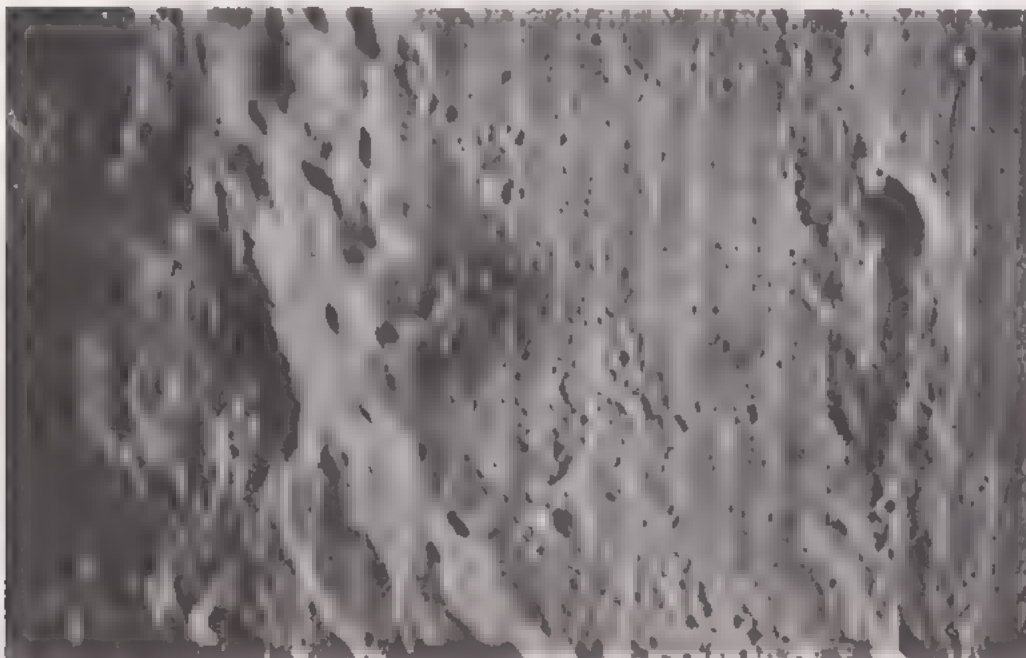


Figure 5.9 Dark halo craters within the lunar crater Alphonsus (120 km in diameter). The evidence for the volcanic origin of these craters is circumstantial but convincing: they are irregularly shaped, rather than circular, located on obvious fissures and fractures, and are surrounded by unusually dark deposits. (Alphonsus is near the centre of the Moon, and can easily be located with binoculars, but the dark halos require a large telescope.)

Promising-looking deposits of this sort have been noted in a few sites around the Moon. Termed **dark halo craters**, they are best seen on the floor of the large crater Alphonsus. There, aprons of dark ejecta surround small craters developed on fissures cutting the floor of the larger crater. They have been described as the lunar counterparts of small terrestrial volcanoes such as Stromboli (Figure 5.9).

5.6.3 Mercury

Only one spacecraft, Mariner 10, has visited Mercury and it imaged only half of the planet. Consequently, we know much less of its volcanic history than we do of the Moon, or even Io, which is more than ten times as distant. Mercury is a **small, dense, body, lacking an atmosphere**. Thus, we would expect it to demonstrate evidence of volcanic activity similar to the Moon's. Mariner 10 images show large areas of 'smooth plains' on Mercury, notably around a large impact basin called Caloris, which are probably equivalents of the lunar mare lavas (Figure 5.10).

5.6.4 Venus

Venus is close in size, mass and density to the Earth, but is much different in having an atmospheric pressure 90 times greater than the Earth's, and a surface temperature of about 460 °C. The crushing atmospheric pressure would tend to limit both explosive pyroclastic activity and the formation of convecting eruption columns.

What effect would the high surface temperature have on effusive activity?

■ The rate of cooling of lava flows would be reduced, so they should flow further.

During 1992–3, while this Course was being prepared, the Magellan spacecraft was still mapping the surface of Venus by imaging radar (as seen in TV programme 4, *Venus unveiled*), so many details of Venus's volcanism remain to be resolved. It is already clear, however, that there are numerous major volcanoes on Venus, rising many kilometres above its surface to form broad shields akin to those of Hawaii (Plate 2.19). Many long lava flows have been identified – some appear to extend nearly 2000 km. There are innumerable small volcanoes, a few hundred metres high. Some of these appear to have been the sources of minor pyroclastic eruptions (Figure 5.11). There are also a number of circular, pancake-like domes (Figure 5.12).

Large areas of the planet are covered by smooth, uncratered expanses of lava, which are evidently relatively young. The *average* age of Venus's surface, in fact, is thought to be only about 500 million years. Given that Venus is so similar to the Earth in size and mass, it is highly probable that volcanic activity continues at the present day. There is no direct evidence of this yet, however.



Figure 5.10 Mariner 10 image of Mercury, showing the Caloris basin, which resembles the Orientale Basin on the lunar far side (cf. Figure 4.11e).



Figure 5.11 Two small Venusan volcanoes (centre) appear to have been built up over basement rocks characterized by geometrically regular fracture patterns. The volcano nearest the centre, which is 3–4 km in diameter, appears to have been the source of the bright pyroclastic deposit extending in a plume which fades out about 10 km from the vent.



Figure 5.12 A puzzling cluster of seven pancake-like domes in the Alpha Regio area of Venus. These domes, which average 25 km in diameter and are about 750 m high, have been compared to extrusions of viscous lavas on Earth. They are, however, much more symmetrical than terrestrial lava domes, and almost an order of magnitude larger

5.6.5 Mars

Much was learned about Mars's geology from the wealth of high-resolution images obtained by spacecraft in the 1970s. It was hoped that much more would be learned by the Mars Observer mission, due to reach Mars in August 1993. Sadly, however, contact was lost with the spacecraft shortly before it arrived.

ITQ 5.6 Before reading any further, note down what sort of volcanism you might expect to find evidence for on Mars. Think in terms of its sources of heat, thermal evolution, and the likelihood of lava and pyroclastic eruptions.

The volcanic processes described in Sections 5.4 and 5.5 should have helped you to formulate reasonable responses to ITQ 5.6. A complicating factor on Mars, however, is that surface conditions, including its atmospheric pressure, appear to have varied greatly since its origin. These changes were reflected in its volcanic history, notably the style of pyroclastic eruptions.

Spacecraft images show that Mars has the largest volcanoes in the Solar System. Olympus Mons (Plate 2.13) and three nearby major volcanoes lie in an area known as the Tharsis volcanic province. Their morphology is reminiscent of the gently sloping Hawaiian volcanoes, but they are much bigger. Olympus Mons has a basal diameter in excess of 600 km, a nested summit caldera complex more than 60 km in diameter and 2 km deep, and rises nearly 25 km above the surrounding plains. High resolution images show myriads of fresh looking flows draping the flanks of Olympus Mons and its neighbour Ascraeus Mons (Figure 5.13), and winding for hundreds of kilometres across the lowland plains.

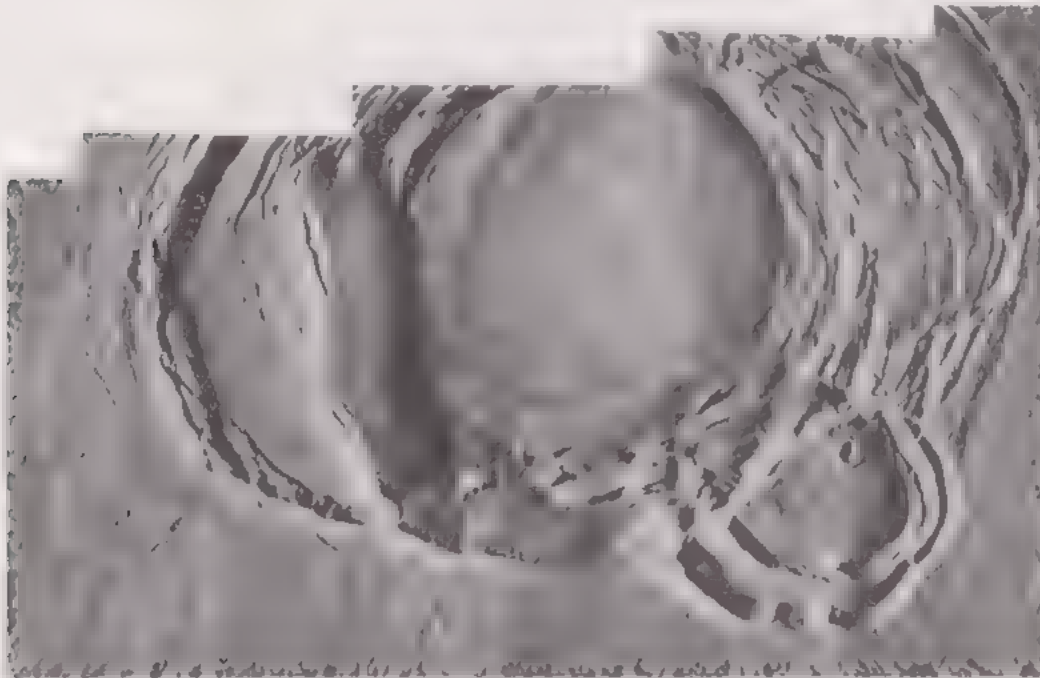


Figure 5.13 The summit caldera complex of Ascraeus Mons, a gently sloping volcano more than 200 km in diameter. Up to eight separate collapse events were involved in the formation of the 45 km diameter crater complex. The floor of the deepest pit lies some 3.5 km below its rim. Numerous small lava flows, up to 1 km wide, are truncated by scarps, showing that subsidence took place later in the evolution of the volcano than lava effusion.

Plate recycling movements do not take place on Mars. Can you suggest how this might account for the fact that even the Earth's largest volcanoes – those in Hawaii – are much smaller than those on Mars?

■ The Hawaiian volcanoes, riding on their lithospheric plate, are rapidly carried away from their thermal sources, thus limiting the maximum size they can achieve. By contrast, a volcano borne above a Martian mantle hot spot remains fixed above it, enabling it to grow indefinitely.

Some small Martian volcanoes rise abruptly out of the surrounding plains, leading to suggestions that they were partially 'drowned' by hugely extensive flood-basalt lavas. These lavas may be up to 4 km thick in the Tharsis area (Figure 5.14).

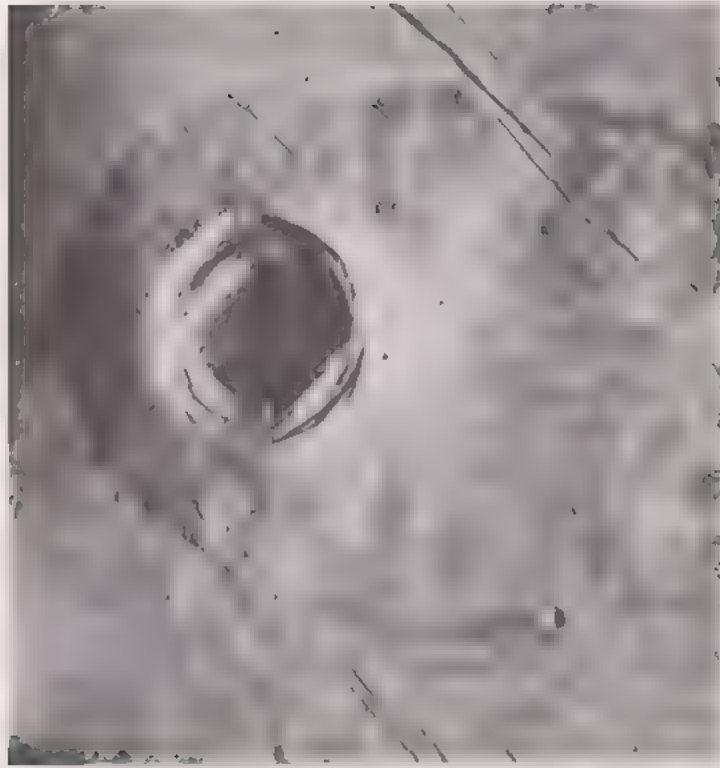


Figure 5.14 Biblis Patera, a small Martian volcano which rises so abruptly above the surrounding lava plains (left side) that, according to the best interpretation, its lower flanks have been entirely drowned by lava. Its caldera is about 55 km in diameter.

Typical flows on the plains are about 10 km wide and 20–30 m thick. Their near-pristine morphologies demonstrate that *some* flows are relatively youthful. With only crater size frequency statistics to go on, however, assigning absolute ages is difficult. Some crater-counters suggest it is almost a billion years since the last eruption; others that the youngest flows may be only a few hundred million years old, **much younger than the youngest flows on the Moon.**

Some Martian volcanoes may be largely of pyroclastic origin. Two lines of evidence support this suggestion, their remarkably flat topographic profiles, consistent with widespread dispersal of pyroclastic fall and flow deposits, and the erosion patterns of the deposits on their flanks, which do not look like those characteristic of lavas (Figure 5.15). Mars's low atmospheric density should allow eruption columns to rise about five times higher than on Earth, at the same mass eruption rate. For comparable eruptions, therefore, pyroclastic deposits could be much more widely dispersed on Mars than on Earth. Whenever things become widely dispersed, however, it becomes more difficult to identify them securely on spacecraft images. Thus, while explosive volcanic activity has probably been important on Mars, it will be difficult to resolve its full extent **without going there.**

Interpretations of spacecraft images suggest that pyroclastic activity was **concentrated in Mars's early volcanic history, and that lava eruptions predominated later.** What conclusion could you draw from this?

■ Most of the volatiles necessary to drive pyroclastic eruptions could have been **lost.** This is consistent with Mars's small size, and tenuous present-day atmosphere.

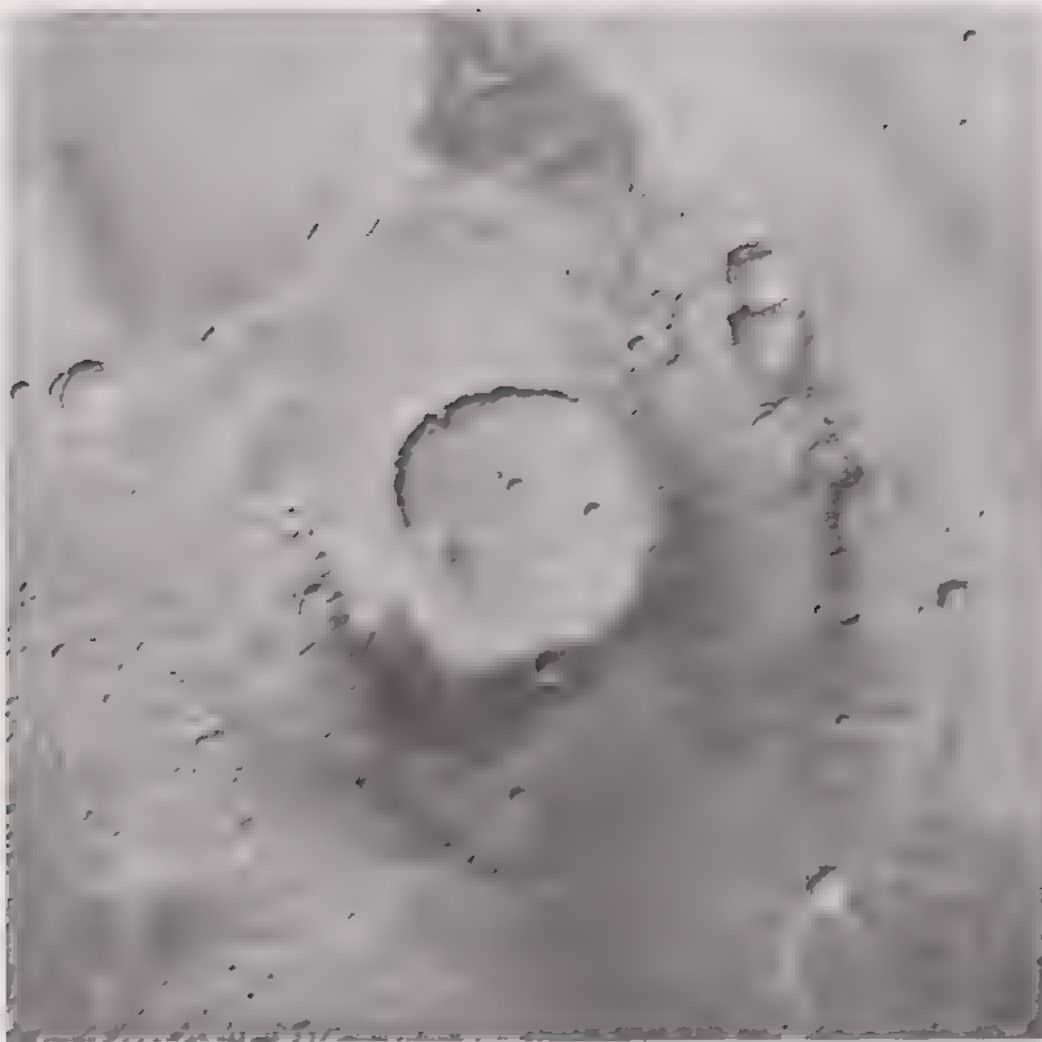


Figure 5.15 Apollinaris Patera, an isolated Martian volcano some 400 km in diameter which has a 70 km diameter summit caldera. Erosion patterns on the flanks of the volcano suggest that they may be composed largely of pyroclastic flow deposits.

5.6.6 Io

Volcanism is so intensely active on Io that it is worthy of special attention. Io is by far the most volcanically active object in the Solar System. Serious scientists have described it as 'a wonderland of physics and chemistry'. Io's density of 3570 kg m^{-3} shows that it should have an iron or iron sulphide core and silicate mantle. As you saw in video sequence 5, Io orbits only 421 600 km away from the colossal mass of Jupiter, and is locked into an elegant series of resonances with the adjacent satellites. These resonances maintain Io in a slightly eccentric orbit, so that Io's tidal bulges oscillate about their mean position, while their heights change according to how close Io is to Jupiter at any moment. Perpetual gravitational kneading keeps Io's interior hot.

This state of affairs had been predicted from theoretical considerations months before images from Voyager 1 confirmed that volcanism was intensely active on Io. In a brilliant paper, three American planetary scientists estimated the power liberated by tidal dissipation to be about $4 \times 10^{13} \text{ W}$, two orders of magnitude larger than that released by radioactive decay (about $5 \times 10^{11} \text{ W}$). Voyager data suggest that the **surface heat flow** on Io is about 1.2 W m^{-2} . By contrast, average surface heat flow on the Earth is 0.08 W m^{-2} , while the Moon's is a mere 0.02 W m^{-2} .

ITQ 5.7 Heat is produced in the Earth at about the same rate as in Io, namely at about $4 \times 10^{13} \text{ W}$. Given this similarity, why does Io's heat production result in so much more obvious and spectacular volcanic manifestations?

Volcanoes on Io

Over 300 volcanoes (not all of them active) have been identified on Io, covering about 5% of its surface area. Most are low, gently sloping shields, less than 100 km in diameter (Plates 2.26–2.28). Radial lava flows snake away from the summit regions, some reaching as much as 700 km from the volcano (Figure 5.16). Many shields have shallow summit calderas ranging from 2–200 km in diameter. A few calderas such as Loki contain features thought to be lava lakes. Others, such as Creidne, have dark lava flows covering the caldera floor, reminiscent of the lava flows filling terrestrial calderas.

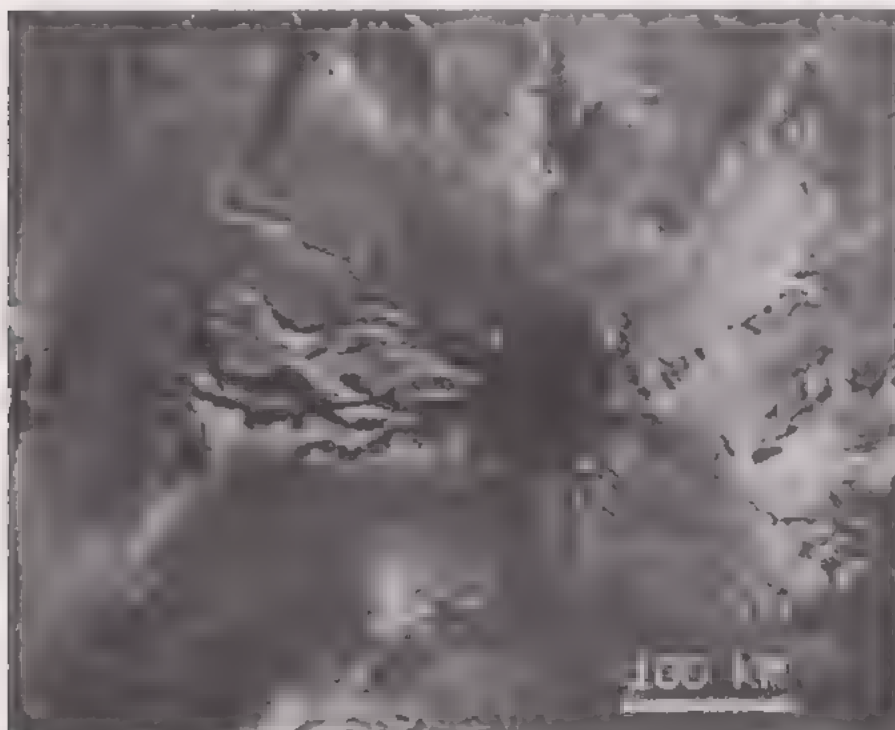


Figure 5.16 Lava flows from Ra Patera on Io extend radially several hundred kilometres from the dark summit caldera. The composition of the lava flows remains undetermined.

Plumes

During the Voyager 1 mission, nine distinct eruption plumes were observed above the surface of Io. They reached heights of between 70 and 280 km, implying eruption speeds between 500 and 1 000 m s⁻¹. Some changed visibly over the course of a few hours, others remained apparently unchanged for days. The largest plume was 1 000 km in diameter, and was fed from a fountain-like vent on a volcano about 35 km across (Plates 2.27 and 2.28). Eight of the plumes were re-observed during the Voyager 2 encounter. In the four months between the two missions, there were considerable changes on Io.

In practical terms, Io lacks an atmosphere. How, then, would you expect the ejecta from its volcanic vents to be distributed?

■ In the absence of an atmosphere, convecting eruption columns cannot be developed. Erupted material should fall back on parabolic trajectories to form sheet-like deposits around the vent areas.

In fact, surface halos (Plate 2.27) around volcanoes such as Pele are good evidence for the accumulation of pyroclastic deposits.

Sulphur or silicates on Io?

All of the above discussions beg the important question: *what is it that is being erupted so prodigiously from Io?* Io's distinctive reddish colour prompted initial speculation that the surface was covered by sulphur, and that the volcanism was entirely sulphurous – the lava flows and active lava lakes were presumed to be made of sulphur, and the active plumes to be driven by sulphur dioxide. This obsession with sulphur was helped by dissemination of so-called 'pizza pictures' of Io, in which image processing techniques were used to colour the satellite in appetizing shades of tomato red and mozzarella yellow.

There are sound reasons for suspecting that sulphur plays an important role in Io's volcanism. Spectroscopic evidence suggests that large amounts of sulphur dioxide and lesser amounts of hydrogen sulphide and carbon dioxide are present in gaseous form above the surface, with a 'frost' of solid sulphur dioxide on the surface. Ionized sulphur has been detected dispersed in a torus (a ring-doughnut shape) along Io's orbit about Jupiter.

A further reason for supposing that sulphur is deeply involved in Io's volcanism is that sulphur dioxide appears to be the most plausible volatile species to drive the eruptions, unlike the Earth, where water is the crucial volatile. When Io was formed, it was so close to Jupiter that it was probably too hot for much water to condense, so it started life fairly dry (Section 2.6).

Can you suggest another reason why Io should not possess much water?

■ Io is a small object, about the same size as the Moon. Thus, gases would be **lost into space, as outlined in Section 1.4.**

The escape speed for Io is about 2.56 km s^{-1} . At Io's likely surface temperature of 100 K, the root mean square speed for water molecules is about 0.4 km s^{-1} (Figure 1.10), while for sulphur dioxide it is only about 0.2 km s^{-1} , so water will be lost at a much faster rate than sulphur dioxide. Moreover, volcanism aids escape by spraying gases outwards – Io has been so active for so long that it is likely that such water as it accreted to start with has been driven off into space by the incessant volcanism.

Thus, sulphur dioxide is probably the volatile component driving Io's volcanism. After being sprayed as vapour hundreds of kilometres into the chill of space, it condenses to the solid form, falling as a snow of tiny particles to the surface, each event forming a thin, even coating. Ultimately, the deposit will become so thick that the lowest layers become warm and mobile enough to be recycled.

So much for the volatile phase. But what of the lavas? Exotic early ideas of lakes of molten sulphur within the calderas of volcanoes such as Loki have been reconsidered in the light of two constraints.

First, sulphur is simply not strong enough to form the 2 km high cliffs that form isolated hills and the walls of calderas. These would simply deform glacier-like under their own weight and fade away, if they were made of sulphur. **The higher the temperature, the more quickly they would deform.**

Second, some thermal phenomena on Io are *too hot* to involve sulphur. One study (using an Earth-based infrared telescope) showed that a 'hot-spot' on Io, probably one of the lava lakes, was about 30 km in diameter, with a 'model' temperature of 873 K. Since the boiling temperature of sulphur in a vacuum is only 715 K, it cannot form the *dominant* magma on Io, although it could form interesting side-shows, as it does on terrestrial volcanoes.

Overall, therefore, it seems most likely that it is silicate volcanism which is so impressively active on Io.

You have now reached the end of our discussion of processes affecting the solid parts of major planetary bodies, though we shall return to some aspects in connection with meteorites in Chapter 8. For the next two chapters, though, we turn to atmospheric processes.

Summary of Section 5.6 and SAQs

- 1 There is abundant evidence for volcanism on the surface of all the larger silicate bodies in the Solar System. It is active at the present day on Earth and Io, and most probably on Venus as well.
- 2 Effusion of basaltic lavas has been the most widespread manifestation of volcanism. Lavas cover 70% of the Earth's surface, and large fractions of the surfaces of the Moon, Mercury, Venus and Mars.
- 3 Earth, Venus and Mars have atmospheric pressures sufficiently large to support convecting eruption columns, drastically modifying the dispersal of pyroclastic ejecta. Pyroclastic flow deposits are extensive on Earth and Mars.
- 4 Io is unique among silicate bodies in that its volcanism is driven by dissipation of tidal energy, and its dominant volatile is probably sulphur dioxide, rather than water. Its lavas, however, are more likely to be silicates, rather than sulphur.

SAQ 5.8 (Objectives 5.2, 5.3 and 5.5) Which of the following gases is the most likely plausible component of Io's eruption plumes, in addition to SO₂, and why? (i) Hydrogen, (ii) helium; (iii) ammonia; (iv) methane; (v) nitrogen; (vi) oxygen; (vii) carbon dioxide; (viii) carbon disulphide.

SAQ 5.9 (Objectives 5.2 and 5.3) Io's abundant heat production has implications for the extent of partial melting in its mantle, and therefore for the compositions of the lavas erupted. Describe in one or two sentences the most likely kind of lava that might be erupted, giving your reasons.

Objectives for Chapter 5

After studying Chapter 5 (and any associated audio, video or TV material), you should be able to:

- 5.1 Give brief definitions of the terms, concepts and principles listed below.
- 5.2 Predict the likely composition of volcanic melts and volatiles arising in silicate and icy bodies, and account for the widespread abundance of basalt lavas in the Solar System.
- 5.3 Predict the varying effects of gravity, atmospheric density and viscosity on volcanic phenomena operating in different planetary contexts.
- 5.4 Outline the factors governing the ascent of convective eruption columns in a planetary atmosphere, and in given circumstances, predict whether pyroclastic fall or flow phenomena would result.
- 5.5 Give brief summaries of the volcanic history of the terrestrial planets and satellites, and the sources of heat driving volcanic activity.
- 5.6 Given spacecraft pictures of planetary surfaces, correctly identify and describe any evidence for volcanic phenomena.

List of scientific terms, concepts and principles used in Chapter 5

Term	Page	Term	Page	Term	Page
aerodynamic drag coefficient	104	eruption rate	100	pyroclastic eruption	99
basalt	96	explosive volcanism	99	pyroclastic flow	105
caldera	105	fire fountain	102	pyroclastic rock	99
chondritic Earth model	95	gas thrust region	103	specific heat capacity	97
compatible elements	96	incompatible elements	96	surface heat flow	115
continental flood basalt	100	isopach	105	terminal fall speed	104
convective ascent region	103	isopleth	105	umbrella region	103
cryovolcanism	96	komatiite	96	viscosity	99
dark halo craters	110	latent heat of fusion	97	yield strength	102
effusive volcanism	100	mid-ocean ridge basalt (MORB)	96		
eruption column	103	partial melting	96		

Chapter 6

Atmospheres of terrestrial planets

Prepared for the Course Team by Charlie Harding

Contents

6.1	Introduction	121
6.2	Exploration	122
	Summary of Section 6.2 and SAQs	124
6.3	Composition of the atmospheres	125
6.4	Atmospheric structure and motion	127
	6.4.1 Effective temperature	127
	6.4.2 Temperature profiles	128
	6.4.3 Cloud formation	131
	6.4.4 Atmospheric motion	134
	Summary of Section 6.4 and SAQs	138
6.5	Atmospheric chemistry	139
	6.5.1 Chemical reactions	139
	6.5.2 Chemistry of the Earth's troposphere	143
	6.5.3 Chemistry of the Earth's stratosphere	147
	6.5.4 Chemistry of the atmosphere of Venus	149
	6.5.5 Chemistry of the atmosphere of Mars	151
	Summary of Section 6.5 and SAQs	153
6.6	Origin and evolution	154
	6.6.1 Origins of the atmospheres: some evidence	154
	6.6.2 Loss of volatiles	156
	6.6.3 Gain of volatiles	157
	6.6.4 Evolution of the atmospheres	159
	6.6.5 The atmosphere of Venus	159
	6.6.6 The atmosphere of Mars	161
	6.6.7 The atmosphere of Earth	162
	Summary of Section 6.6 and SAQs	165
	Objectives for Chapter 6	166

6.1 Introduction

In this chapter we turn from our examination of the solid planetary bodies to explore the atmospheres that envelop three of the terrestrial planets: Venus, Earth and Mars. Mercury has an extremely tenuous atmosphere, about which we shall say very little. On Earth, the atmosphere is of special interest to us because of its role in sustaining life. Biological activity has been the main influence on the composition of the Earth's atmosphere, making it unique amongst the atmospheres in the Solar System. Life, and especially human activities, continue to influence the Earth's atmosphere, so that there is much speculation about the climatic changes that these influences may cause.

As the description of the atmospheres of the terrestrial planets progresses through this chapter, so it becomes increasingly speculative, ending on the controversial topic of the origin of atmospheres and their evolution.

We begin on a more certain note: the exploration of atmospheres (Section 6.2). Since the early 1960s our knowledge of the planets has expanded enormously. Even before the space age, remote sensing by spectroscopic methods enabled us to determine some components of planetary atmospheres. Nevertheless, even as late as the 1950s, little was known of the composition of the atmospheres of other planets (see Figure 6.1), and even less about their physical properties, such as temperature and pressure.



Figure 6.1 Dan Dare lands on Venus.

With a knowledge of the composition, physical properties and cloud layers of the atmospheres we are able to compare them (Section 6.3). First inspection suggests that they have little in common, but a more careful comparison hints at a common origin.

On Earth we are all familiar with wind and weather. Not surprisingly, winds occur on other planets, and in Section 6.4 and video sequence 7 we examine the causes of atmospheric motion and compare their effects on different planets. The major contributor to atmospheric motion is the energy from the Sun. The interaction of solar radiation with the atmosphere also causes layering, giving atmospheres a vertical structure. In the case of the Earth, this layering is another factor that sets its atmosphere apart from those of other planets.

A further manifestation of the interaction of solar radiation with planetary atmospheres is chemical change. Radiation at the high-frequency end of the solar spectrum disrupts molecules. Consequently, we might expect atmospheric composition to change with time. For Earth and, as far as we know, for the other terrestrial planets, it remains roughly constant, the result of chemical cycles that regenerate molecules and restore composition. On Earth, life has a major effect in maintaining this composition. We review some chemical cycles that determine atmospheric composition in Section 6.5.

In spite of short-term chemical stability, there exists much evidence for long-term chemical evolution of our atmosphere. This is the subject of Section 6.6, in which we also consider, on the basis of much thinner evidence, the evolution of the atmospheres of our neighbouring planets.

Many of the ideas introduced in this chapter are of general applicability; they are relevant to other atmospheres. In Chapter 7 we consider the giant planets, for which our knowledge of the outer atmospheres is more soundly based than that of the interiors, which remain hidden from us. So in the next chapter you may expect to find further applications of the principles that are discussed here.

6.2 Exploration

Because we are immersed in the Earth's atmosphere, our knowledge of at least its lower region has developed with the recognition that it is composed of different gases. This development dates from the eighteenth century, with the discovery of the **diatomic molecules**, oxygen (O_2) and nitrogen (N_2). Knowledge of this composition was extended by the identification in the last decade of the nineteenth century of the **noble gases**, listed in Table 6.1. The discovery of these unreactive gases was made, not surprisingly, by using electronic spectroscopy, a technique devised some thirty years earlier (Book 1, Subsection 1.3.2). The identification of the other minor components of the Earth's atmosphere depended on the development of other techniques, which have also proved useful in studying the atmospheres of other planets.

These methods can be divided into two types. Techniques that can be used remotely are particularly effective in studying distant planets; techniques that require direct access to material under investigation are limited to missions that enter atmospheres, although they often provide more detailed information than remote sensing methods.

Apart from electronic spectroscopy, what other techniques would you expect to be useful for remote sensing of molecules?

■ In addition to electronic spectroscopy, **vibrational spectroscopy** and **rotational spectroscopy** can be used, as in the study of interstellar clouds (see Book 1, Subsection 5.2.5).

For planetary atmospheres, where temperatures are high relative to those of the interstellar medium, vibrational spectroscopy is of more use, especially where there is some infrared source that enables the absorption spectrum of an atmosphere to be studied. Thermal emission from the surfaces of the planets provides such sources, because their continuous black-body radiation spectra peak in the infrared.

Vibrational absorption spectroscopy, as you learned in Book 1, arises from the interaction of infrared radiation with molecules. We envisage electromagnetic radiation as the propagation of oscillating electric and magnetic fields (Book 1, Subsection 1.2.4). For the electric field to interact with a vibrating molecule two conditions must be met. First, the molecule must have an uneven distribution of electric charge, with one end positive and the other negative, as illustrated in Figure 6.2. We call this distribution of charge an **electric dipole**, and it arises in molecules when one atom has a greater tendency to attract electrons to itself than some neighbouring atom. The second condition requires the frequency of the electromagnetic radiation to be the same as the frequency of vibration of the molecule: during the vibration of the molecule, the atoms are displaced as the bond alternately lengthens and shortens, and absorption occurs when the change in value of the electric dipole is at the same frequency as the field oscillations (Figure 6.2).

A diatomic molecule consists of two atoms bound together.

Table 6.1 The noble gases

Element	Symbol
helium	He
neon	Ne
argon	Ar
krypton	Kr
xenon	Xe
radon	Rn

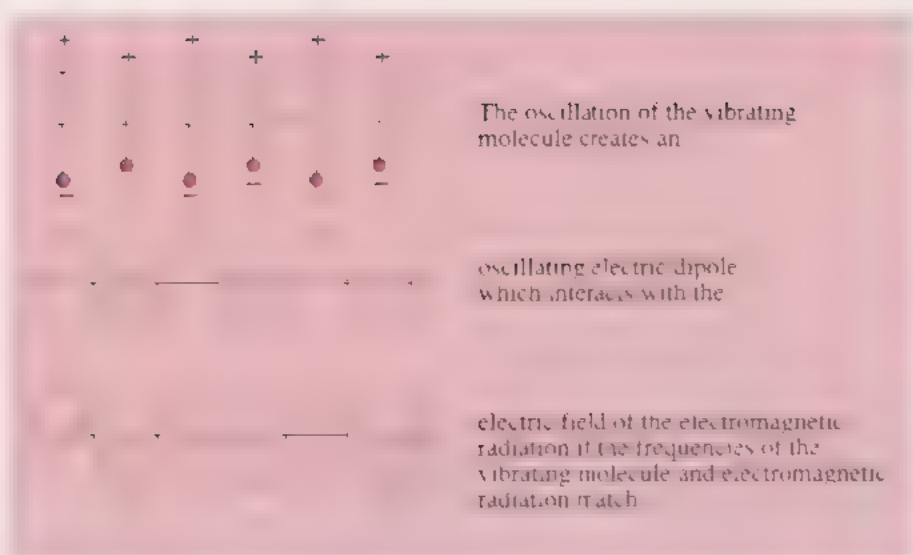


Figure 6.2 The oscillating electric field of electromagnetic radiation interacts with the electric dipole of a diatomic molecule, if the frequencies match

Amongst diatomic molecules, which type would you expect *not* to interact with infrared radiation via their vibration?

■ The molecules that will not interact are those without an electric dipole, i.e. those containing atoms that have an equal tendency to attract electrons. Of course, if the two atoms are identical this condition applies.

We call molecules that consist of identical atoms **homonuclear molecules**, and those containing different types of atom **heteronuclear molecules**.

So we conclude that homonuclear diatomic molecules, such as H_2 and N_2 , will not interact with electromagnetic radiation during their vibration and so will not have a vibrational spectrum. As a general rule, for a diatomic molecule to absorb radiation through its vibration, it must possess an electric dipole. Rules such as this, which determine the spectroscopic transitions that can be observed, are called **selection rules**. Although this rule restricts the diatomic molecules that can be detected by vibrational spectroscopy, all molecules with more than two atoms have vibrational spectra.

Direct sensing methods can provide more information, and two, in particular, have been used in the study of atmospheres. These are gas chromatography and mass spectrometry. For our purposes, it is more important to take note of the information that is obtained by these techniques than the details of their application.

In principle, **gas chromatography** is a means of separating a complex mixture of gases into its chemically distinct components. This is achieved quite simply by pumping a small sample of the mixture along a tube that is coated with a liquid layer. Gases that tend to dissolve in the liquid take longer to emerge from the tube than those that are insoluble. Provided the instrument has been calibrated with authentic samples of the component gases their identification is often possible, but without such a calibration it is difficult to identify the components of a mixture. In planetary exploration it is therefore important to anticipate the results of analyses by this technique, or to couple it with a method of identifying the components independently. One such method is **mass spectrometry**, which you will learn more about in video sequence 8. It provides a way of measuring the masses of molecules, and of fragments of molecules if they break up on ionization, as they often do. These masses are usually expressed as **relative molecular mass (RMM)**.

The RMM is defined as the mass of the molecule relative to the mass of an atom of an isotope of carbon, ^{12}C , which is given the value of 12 on the RMM scale:

$$\text{RMM (molecule)} = 12 \times \frac{\text{mass of molecule}}{\text{mass of } ^{12}C \text{ atom}}$$

Recall from Subsection 5.2.5 of Book 1 that homonuclear diatomic molecules do have extremely weak vibrational and rotational spectra, but these can be ignored for our purpose here.

The use of the RMM scale conveniently avoids the powers of ten involved in the masses of individual molecules. For example, the mass of a molecule of carbon dioxide, $^{12}\text{C}^{16}\text{O}_2$, is 7.31×10^{-26} kg, whereas its RMM is 43.99. RMM values can be estimated approximately by summing the mass numbers of the constituent atoms.

Can you envisage any problems in the application of mass spectrometry, e.g. in identifying molecules such as CO and N_2 composed of their most abundant isotopes: ^{12}C , ^{14}N and ^{16}O ?

■ These two molecules have approximately the same RMM of 28, and so will not be distinguished by mass spectrometry except at high resolution.

High-resolution mass spectrometers are large and expensive, and it has not been practical to despatch them to distant planets.

Mass spectrometry has, however, a distinct advantage over many other analytical methods. As it measures RMMs, it enables us to distinguish between isotopic variants, e.g. between $^{12}\text{C}^{16}\text{O}$ and $^{13}\text{C}^{16}\text{O}$, the carbon monoxide molecules containing the isotopes ^{12}C and ^{13}C , respectively. This facility has provided results that have been of great value as evidence for the origin of planetary atmospheres.

The application of these techniques in fly-by and lander missions to Mars and Venus has provided increasingly detailed descriptions of these planets since the first Mariner fly-by of Venus in 1962. The Soviet Venera and Vega missions to Venus over the period 1965 to 1986 have given much information about the physical and chemical properties of the atmosphere of Venus, augmented by the American Pioneer missions of 1978. A series of American missions to Mars dating from the 1965 Mariner spacecraft to the Viking landers of 1976, and the Soviet Phobos mission, have provided a detailed description of the Martian atmosphere.

The picture that emerges of these three terrestrial planets suggests bodies with atmospheres as contrasting as their appearance from a distance, shown in Plates 2.29, 2.30 and 2.31. Earth has an atmosphere of complex structure with much seasonal activity. Clouds of water droplets and ice (H_2O) crystals are extremely variable and cover typically about half of the planet. Venus has an atmosphere that is about 100 times as massive as that of the Earth, it shows no seasonal change, and it contains an unbroken layer of dense cloud that obscures the planet's surface. The atmosphere of Mars is tenuous and undergoes extreme **diurnal** (day–night) and seasonal **variations**, with occasional massive dust storms.

Summary of Section 6.2 and SAQs

- 1 Remote sensing of planetary atmospheres provides information about their composition. Infrared (vibrational) spectroscopy is applicable to all atmospheric gases, except for those diatomic molecules that possess no electric dipole (i.e. homonuclear diatomic molecules). O_2 , N_2 , Ar
- 2 Analysis by direct access of planetary atmospheres has successfully employed two major techniques, gas chromatography and mass spectrometry. gas chromatography is used to separate complex gaseous mixtures; mass spectrometry provides the relative molecular mass of a substance, and can also give information about isotopic composition.

SAQ 6.1 (Objectives 6.1 and 6.2) Comment on the use of mass spectrometry and of gas chromatography in analysing a mixture of the following substances: CO_2 , O_2 , N_2 , C_2H_2 and C_2H_4 . Assume that these molecules are composed of the most abundant isotopes (^{12}C , ^{16}O , ^{14}N and ^1H).

SAQ 6.2 (Objectives 6.1 and 6.2) Which of the following substances would it be possible to detect remotely by infrared spectroscopy CO , H_2 and HCl ?

6.3 Composition of the atmospheres

The atmospheres of the planets show great diversity. Although we group Venus, Earth and Mars together as terrestrial objects, their atmospheres appear, at least at first inspection, to display more differences than similarities. The major constituents of these atmospheres are listed in Table 6.2 and compared in Figure 6.3. The interaction of solar radiation with an atmosphere leads to the production of reactive species, with ensuing chemical reactions. Consequently, the composition of the atmospheres changes with altitude. The compositions given here are those at the surface, where the atmospheres are most dense.

Table 6.2 The major components of the atmospheres of Venus, Earth and Mars at the surface

Venus ^a		Earth		Mars	
Gas	Volume ratio ^b	Gas	Volume ratio ^b	Gas	Volume ratio ^b
CO_2	0.965	N_2	0.781	CO_2	0.953
N_2	3.5×10^{-2}	O_2	0.209	N_2	2.7×10^{-2}
SO_2	1.5×10^{-4}	H_2O^c	≤ 0.04	Ar	1.6×10^{-2}
H_2O	1×10^{-4}	Ar	9.3×10^{-3}	O_2	1.3×10^{-3}
Ar	7×10^{-5}	CO_2	3.4×10^{-4}	CO	7×10^{-4}
H_2	$< 2.5 \times 10^{-5}$	Ne	1.8×10^{-5}	H_2O	3×10^{-4}
CO	2×10^{-5}				

^a Some of these values are measured at an altitude of 22 km because surface data are unreliable.

^b The volume ratio is the fraction by *number* of the atoms or molecules present. Chemists often refer to this as the mole fraction. It is also called the volume mixing ratio by some atmospheric scientists. When multiplied by the atmospheric pressure, it gives a quantity called the partial pressure (Section 6.4.3), which may be envisaged as the fractional contribution of a component to the total pressure.

^c The H_2O in the Earth's atmosphere is highly variable!

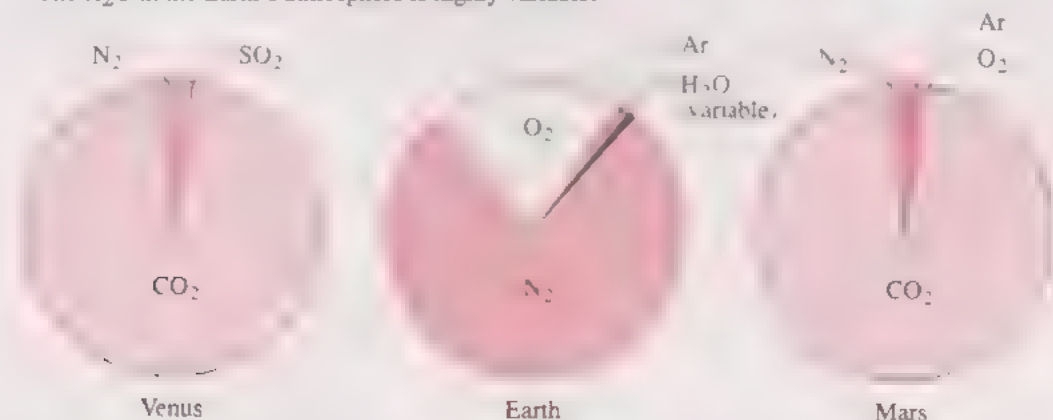


Figure 6.3 The major components of the atmospheres of Venus, Earth and Mars, as measured at the surfaces of the planets. The area of each slice of the pie chart is proportional to the volume ratio of the substance shown.

Column mass was defined in terms of the solar nebula in Subsection 2.3.2.

As you might expect, for a given column mass, pressure increases as gravity increases.

The amounts of atmosphere on these planets also differ greatly. To compare these amounts we use a quantity called column mass, which you met in Subsection 2.3.2. The pressure at the surface of a planet is not a good measure of the amount of atmosphere because this pressure depends on the gravitational field of the planet. **Column mass**, which is given the symbol m_c , is the mass of atmosphere that sits above 1 m^2 of a planet's surface. It may be estimated from measurements of the atmospheric pressure, P , at the surface and the gravitational acceleration, g , at the surface, using the expression

$$m_c = \frac{P}{g} \quad (6.1)$$

For the Earth the value is about 10^4 kg m^{-2} .

IIQ 6.1 Estimate the values of the column mass for Venus and Mars, and compare the column masses of the three terrestrial planets. For Venus: $P = 92 \text{ bar}$; $g = 8.90 \text{ m s}^{-2}$. For Mars: $P = 6.3 \times 10^{-3} \text{ bar}$; $g = 3.72 \text{ m s}^{-2}$. (Remember that $1 \text{ bar} = 10^5 \text{ Pa}$, and that $1 \text{ Pa} = 1 \text{ N m}^{-2} = 1 \text{ kg m}^{-1} \text{ s}^{-2}$.)

Table 6.3 also gives the pressures and average temperatures at the surfaces of the three planets, which also differ markedly. In atmospheres of such different pressures and compositions, it is not surprising to find different clouds.

Table 6.3 The surface temperature T_s , pressure P_s , and column mass m_c of the atmospheres of Venus, Earth and Mars

Planet	T_s/K	P_s/bar	$m_c/\text{kg m}^{-2}$
Venus	732	92.1	1.03×10^6
Earth	288	1.01	1.03×10^4
Mars	223	6.3×10^{-3}	1.69×10^2

On Venus, the cloud layer is thick, extending from about 45 to 65 km above the surface, and extensive, covering the entire planet. The clouds consist largely of particles about $2 \mu\text{m}$ in diameter, whose properties closely match those of droplets of 75% H_2SO_4 (sulphuric acid) and 25% H_2O . Additionally, the clouds contain larger solid crystals of chlorine-containing substances, thought to be FeCl_2 and Al_2Cl_6 , as well as phosphorus compounds, such as H_3PO_4 .

On Earth, we are familiar with clouds of H_2O , which may exist as liquid droplets or as particles of snow or ice. These clouds occur up to altitudes of about 50 km.

On Mars, the major clouds are composed of dust (mostly silicate particles) raised by wind. These storms are quite variable. In addition, clouds of solid particles of H_2O and CO_2 occur, as shown in Plate 2.31, although the total average cloud cover is small.

On each of these planets the clouds reflect solar radiation, and this strongly affects the heating of the planet. The total fraction of solar radiation that is reflected by a planet is called the **albedo**, a . The total fraction absorbed is simply $(1 - a)$. Surface reflection also plays a part, but the influence of cloud cover can be seen by a comparison of the albedos in Table 6.4. Later, when we have considered how temperature and pressure vary with altitude, we shall return to the question of cloud formation.

On Earth, the physical conditions that determine cloud formation and precipitation ensure that most of the planet's surface and atmospheric H_2O exists in the oceans, which act as a reservoir, whereas on Venus water cannot exist as liquid because of the high surface temperature. The Earth's polar ice-caps form another part of the H_2O reservoir. The ice caps observed on Mars (Plate 2.32) are composed of both H_2O and CO_2 . On Earth, CO_2 is largely combined with other

Table 6.4 The albedos of Venus, Earth and Mars

Planet	Albedo a
Venus	0.77
Earth	0.23
Mars	0.14

elements in carbonate deposits, especially as CaCO_3 in sedimentary rocks, whereas on Venus it is the main atmospheric constituent.

So substances that exist as gases on one planet may on other planets be condensed as liquids or solids or may occur within minerals. For some purposes of comparison, it is more valid to consider the total (estimated) inventory of potentially volatile substances by, for example, taking account of all the CO_2 in carbonate rocks and the condensed H_2O on Earth. When this is done we find some striking similarities in the major **volatile inventories**. Each of the three planets contains, for example, about 10^{-4} kg of CO_2 per kilogram of planet. These volatile inventories provide a much better basis for comparison when we come to consider the origins of atmospheres.

Note that 'volatile' is itself a word of relative meaning, it is used here in a somewhat different sense from that in Chapters 2 and 3, where we were concerned with condensation within the solar nebula. Here it has a similar meaning to its use in volcanology in Section 5.5.

6.4 Atmospheric structure and motion

On Earth, it is common experience that temperature decreases at higher altitudes, at least to the tops of the highest mountains. This variation of temperature with altitude is an aspect of **atmospheric structure**, and is determined largely by the absorption of solar radiation by a planet's atmosphere and surface. Together with pressure, which also decreases with altitude, temperature determines the formation of clouds in the atmosphere. The heating of the atmosphere and surface of a planet by solar radiation is also the cause of atmospheric motion.

6.4.1 Effective temperature

The temperature of the surface and atmosphere of a planet is determined by the balance between the energy that is absorbed and the energy that is emitted. For the terrestrial planets, the source of most of the energy reaching the atmosphere and the surface is the Sun; relatively little energy comes from the planet's interior. The solar flux density at the top of the Earth's atmosphere is about $1.38 \times 10^3 \text{ W m}^{-2}$. Some of this is reflected back to space, some is absorbed by the atmosphere, and the rest reaches the surface, where it is either reflected or absorbed. The absorbed radiation heats the surface, which then re-radiates this energy, mainly in the infrared region.

On the assumption that a planet undergoes no net heating or cooling in the short term, it is possible to estimate the temperature necessary for a planet to re-radiate all of the energy absorbed by the atmosphere and the surface. This temperature, called the **effective temperature**, T_e , is defined as follows:

$$T_e = \left[\frac{L}{4\pi R^2 \sigma} \right]^{1/4} \quad (6.2)$$

Here L is the total power radiated by the planet, R is the radius of the planet (its surface area is $4\pi R^2$, and radiation is emitted from the whole surface, Figure 6.4) and σ is Stefan's constant, $5.67 \times 10^{-8} \text{ W m}^{-2} \text{ K}^{-4}$. Notice that Equation 6.2 is similar to the expression for the luminosity of a star in terms of its surface temperature, in Book 1, Subsection 2.3.3.

T_e is an alternative way of expressing the total power L radiated by a spherical body. This relationship applies regardless of the form of the spectrum of the emitted radiation, although for the planets the spectrum is close to that of a black body at the effective temperature T_e .

It is instructive to estimate T_e for the terrestrial planets and to compare these temperatures with those at the planets' surfaces. In order to estimate T_e , the power lost by radiation must be equated with that absorbed from solar radiation. The Sun's radiation arrives from one direction, and so a planet is heated over

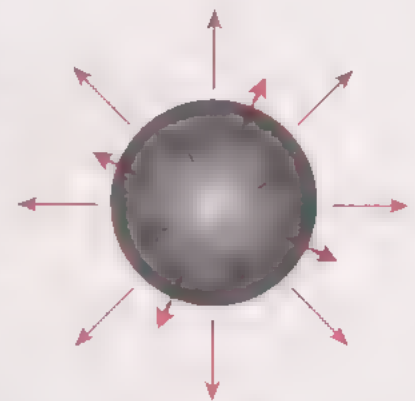


Figure 6.4 The emission of radiation occurs from the whole surface of a planet, which is an area of $4\pi R^2$ for a planet of radius R .

only half of its surface at any time. So the planet casts a disc-shaped shadow of area πR^2 , where R is the radius of the planet (Figure 6.5). The power absorbed depends on this area, and also on the solar flux density at the distance of the planet from the Sun. A third factor that determines the power absorbed by a planet is the albedo. A planet that has a light-coloured surface or extensive cloud cover reflects much of the Sun's radiation, whereas a dark surface absorbs radiation. A planet with a high albedo thus absorbs little of the Sun's radiation. With an average cloud cover of about 50%, the Earth has an albedo of 0.23. By equating the power L radiated by the Earth with the solar power absorbed, which can be readily estimated independently, an effective temperature of 255 K is estimated from Equation 6.2.

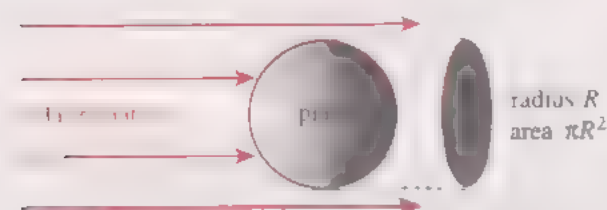


Figure 6.2 The absorption of solar radiation: a planet presents a disc of area πR^2 to the Sun's radiation, and casts a shadow of this area.

The effective temperatures of Venus and Mars are 227 K and 217 K, respectively

□ How do these compare with the value of T_e for the Earth?

■ Both are significantly lower.

□ Why are they lower?

■ Mars is farther from the Sun. Venus, although nearer to the Sun, has a much higher albedo (Table 6.4) because of its total cloud cover.

Notice that this estimate of effective temperature is based entirely on the Sun as a source of energy, and ignores heat sources internal to a planet. So, how much higher would the effective temperature be if heat from the planetary interior were taken into account?

ITQ 6.2 The solar flux density at the radius of the Earth's orbit is $1.38 \times 10^3 \text{ W m}^{-2}$. How does the total solar flux that is absorbed by the Earth compare as a heat source with the total flux derived from geothermal heat flow, given in Subsection 3.5.2?

6.4.2 Temperature profiles

Effective temperature is defined by Equation 6.2 in terms of the power radiated by a planet, and so does not necessarily correspond to an average temperature on the surface of a planet. Temperatures at the surfaces and at altitudes in the atmospheres have been recorded for Venus, Earth and Mars. Of course, these temperatures also vary with latitude, time of day and the seasons. To single out the variation with altitude, these other effects are averaged. These variations of average T with altitude are shown in Figure 6.6.

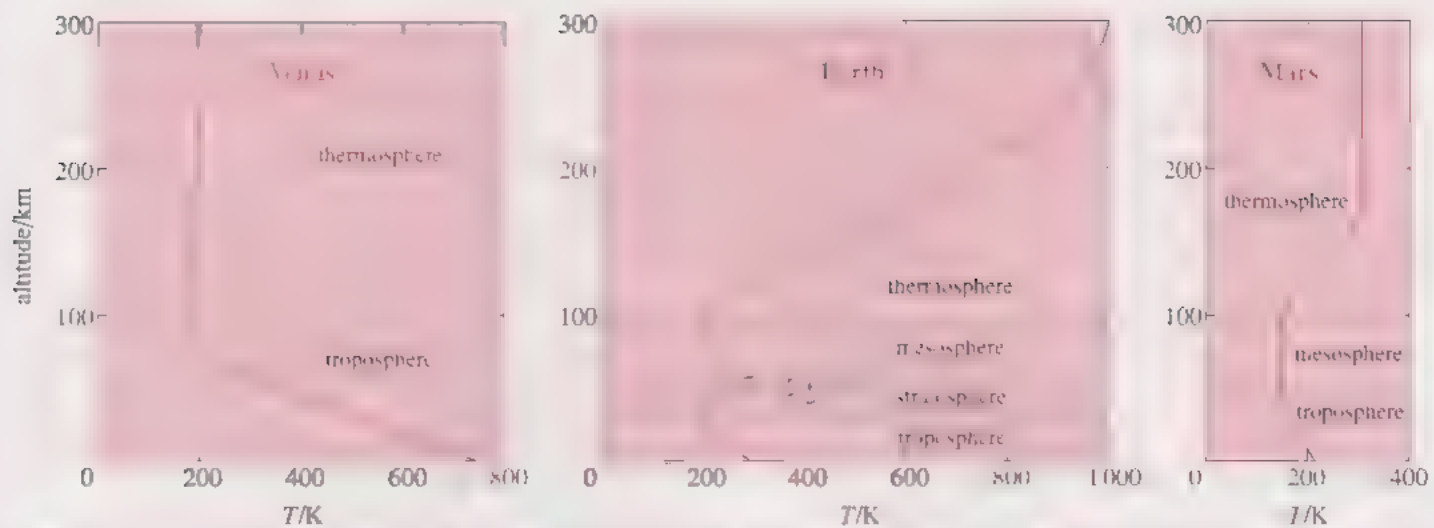


Figure 6.6 Atmospheric structure: the variation of temperature with altitude for Venus, Earth and Mars. The atmospheres are divided into layers corresponding to decreasing or increasing temperature.

Consider the average surface temperatures of the planets, and compare these with the effective temperatures given in Subsection 6.4.1. The average surface temperature of Mars is 223 K, almost exactly equal to its effective temperature.

How do $T(\text{surface})$ and T_e compare for Earth and for Venus?

■ For Earth, the surface temperature is 288 K, which is 33 K higher than T_e , whereas the average surface temperature of Venus is 732 K, about 500 K higher than T_e .

The cause of these high surface temperatures is the so-called **greenhouse effect**, which arises in the following way. The surface of a planet is heated by solar radiation, which has a spectrum in which the energy is concentrated in the visible region. This spectrum is approximately that of a black body emitting at a temperature of 5770 K, the Sun's photospheric temperature. The hot surface of the planet loses heat mainly by radiation from the surface.

11Q 6.3 If the Earth, with a surface temperature of 288 K, radiates as a black body, at what wavelength is the maximum energy of the emission spectrum predicted to be by Wein's displacement law, $\lambda_{\text{peak}} = (2.90 \times 10^{-3} \text{ m K } T)^{-1}$? In which region of the spectrum is this radiation most intense (see Book 1, Figure 1.23)?

The black-body spectra at these temperatures, 288 K and 5770 K, are shown in the upper part of Figure 6.7. The lower half of the figure shows the absorption spectrum of the Earth's atmosphere. The atmosphere is quite transparent in the visible region. However, in the infrared region, electromagnetic radiation is absorbed by molecules which, as a result, increase their vibrational energy (Book 1, Subsection 5.2.5). During subsequent collisions between molecules, this internal vibrational energy is converted, albeit slowly, through chemical reactions and by energy transfer in collisions, into kinetic energy of the molecules. So solar (visible) radiation is transmitted by the atmosphere and heats the surface of the planet. The surface radiates at infrared wavelengths, and this energy is trapped by absorption in the atmosphere, thereby raising its temperature. This heating of the lower atmosphere (and surface) of the planet is the **greenhouse effect**.

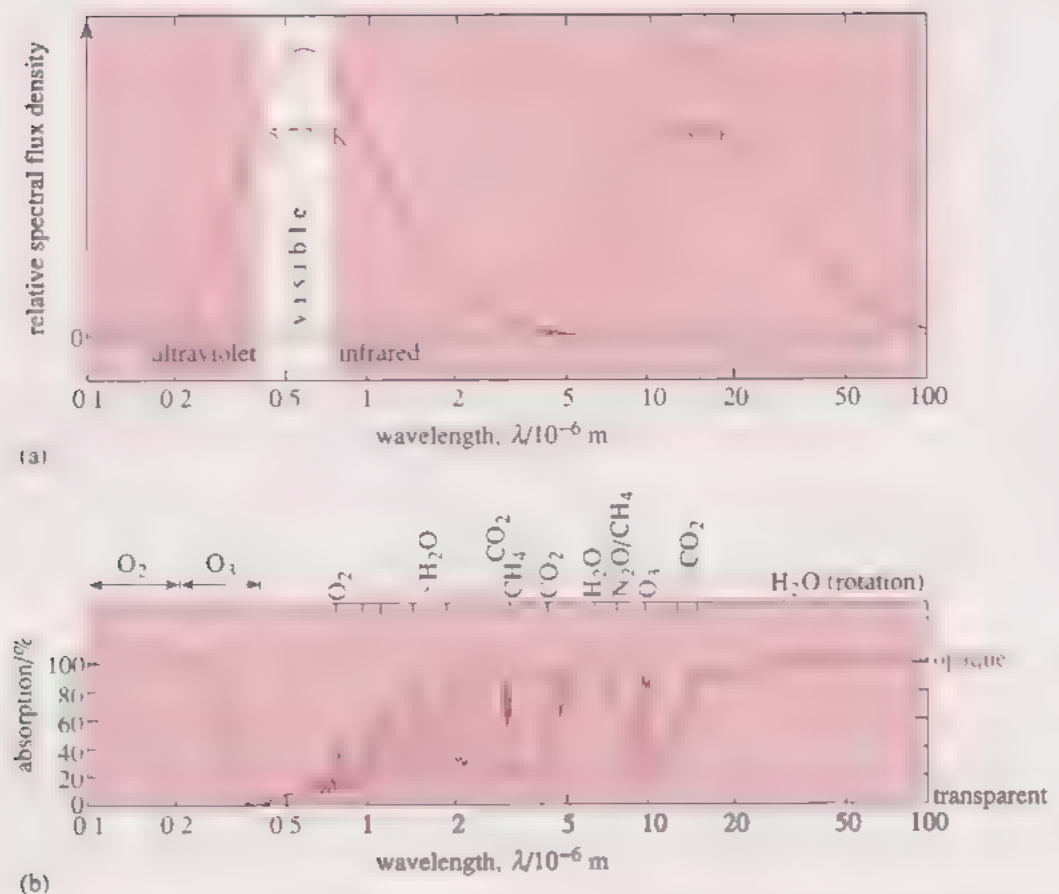


Figure 6.7 (a) Spectra of black-body sources at the temperatures of the Sun's photosphere (5770 K) and the Earth's surface (288 K). The vertical scales for the two spectra are not the same; the Sun's radiation is much more intense than that of the Earth. (b) The absorption spectrum of the Earth's atmosphere; the wavelengths at which some atmospheric gases absorb energy are indicated.

Notice that most of the absorption at the most intense region of the 288 K spectrum is caused by CO_2 and H_2O , minor components of the Earth's atmosphere.

Why do the major components of the Earth's atmosphere, N_2 and O_2 , not contribute to the greenhouse effect?

■ N_2 and O_2 are homonuclear diatomic molecules and so do not possess an electric dipole. So they do not absorb infrared radiation strongly through changes in their vibrational energy (Section 6.2).

The high surface temperature of Venus can be attributed to the absorption of surface radiation by CO_2 . With a much more dense atmosphere than that of Earth, Venus has a large greenhouse effect, causing its surface temperature to be about 500 K greater than T_e .

Currently there is much concern about the environmental effect of raising the level of CO_2 in the Earth's atmosphere, and the effect of this on global warming (Subsection 6.6.7). Nevertheless, it is to the natural greenhouse effect that we owe our existence, as life would be very different, if it existed at all, on a planet that was on average 33 K colder than Earth.

Figure 6.6, which shows the vertical temperature variation of the atmospheres, indicates that several similarities exist between the three planets. In the lowest region of each atmosphere, the temperature drops with increasing altitude. This region is called the **troposphere**, meaning the region of mixing. The rapid decrease of temperature with altitude, which is due to heating of the surface, causes convection, which leads to vertical mixing.

The uppermost region of each atmosphere, called the **thermosphere**, is characterized by an increase of temperature with altitude; this increase shows a

large diurnal variation. Heating of the thermosphere results from the absorption of high frequency (ultraviolet, UV) solar radiation, which causes ionization of gases, mainly O_2 for the Earth, and CO_2 for Mars and Venus.

Eventually much of this energy is converted via recombination of ions and electrons, and by chemical reactions, into vibrational and kinetic energy of the constituents of the thermosphere. For this energy to be radiated, the vibrationally excited molecules must relax to a low energy state with the emission of a photon of infrared radiation, the reverse of the absorption process described in Section 6.2. For the components of the Earth's thermosphere, such as the homonuclear molecule O_2 , this process is very inefficient. So although very little of the solar radiation is absorbed in the thermosphere, it is sufficient to generate a high temperature because the loss of this energy is highly inefficient. By contrast, the thermospheres of Venus and Mars, which consist mainly of CO and CO_2 , are able to radiate efficiently, and so we do not observe temperatures as high as those in the Earth's thermosphere, as shown in Figure 6.6.

Between the Earth's troposphere and thermosphere, at about 20 to 90 km, there is a region in which the temperature first rises and then falls. The lower part of this region (20–50 km) is called the **stratosphere**, and the upper part the **mesosphere**. Evidently there exists some mechanism for absorbing energy, which is responsible for the temperature rise in the stratosphere. The substance responsible for this absorption is **ozone**, O_3 . Although ozone occurs throughout the atmosphere, it reaches its highest concentration in the stratosphere, as Figure 6.8 shows. It is this peak in concentration of O_3 that is called the ozone layer. In fact, the concentration of O_3 is highest in a band between 20 and 30 km altitude, although it is only a minor atmospheric component. If it were all collected and compressed to 1 atmosphere pressure at the Earth's surface, it would form a layer only 3 mm thick.

As solar radiation enters the Earth's atmosphere, it is absorbed selectively by various components of the atmosphere (Figure 6.7). Ionization of molecules in the thermosphere is responsible for the removal of radiation of wavelengths shorter than about 100 nm. Molecular oxygen (O_2) absorbs radiation of $\lambda < 230$ nm (Figure 6.9), and is split into oxygen atoms by radiation of $\lambda < 140$ nm. Ozone absorbs radiation in the wavelength range from 200 to 350 nm, radiation that is transmitted by the thermosphere and largely absorbed in the stratosphere. At wavelengths longer than 230 nm, the only component of the atmosphere that absorbs solar UV radiation significantly is ozone. In spite of its low concentration, it is highly effective at absorbing solar radiation of $\lambda < 350$ nm (Figure 6.9), energy that is responsible for the heating of the stratosphere. Thus, ozone prevents this part of the solar spectrum (230–350 nm) from reaching the Earth's surface. This region of the spectrum is highly damaging to life, causing genetic damage when absorbed by DNA. So life is protected by the ozone layer. The processes by which ozone is produced and destroyed in the stratosphere, and the details of the absorption process, are considered in Subsections 6.5.2 and 6.5.3.

6.4.3 Cloud formation

On each of the terrestrial planets we observe clouds (see Section 6.2); on Earth we are familiar with clouds formed of water. Although it is difficult to forecast when it will rain, it is possible to predict in a model system the conditions when water will exist as vapour, liquid, ice or any combination of these. The two factors that determine the phase in which water will exist are temperature and pressure. To be more precise, it is the **partial pressure** that determines the phase; the partial pressure of H_2O is equal to the total pressure of the atmosphere multiplied by that fraction of the molecules in the gas that are H_2O molecules.

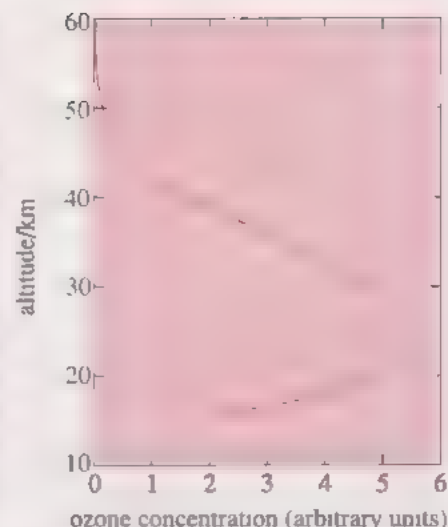


Figure 6.8 The variation of ozone concentration with altitude in the Earth's atmosphere. There is some variation with latitude, but over the whole of the Earth ozone occurs mainly between 10 and 50 km altitude

In reading about ozone, you may come across Dobson units. One Dobson unit (D.U.) corresponds to 0.01 mm thickness of O_3 under these conditions. So the 'ozone layer' is about 300 D.U.)

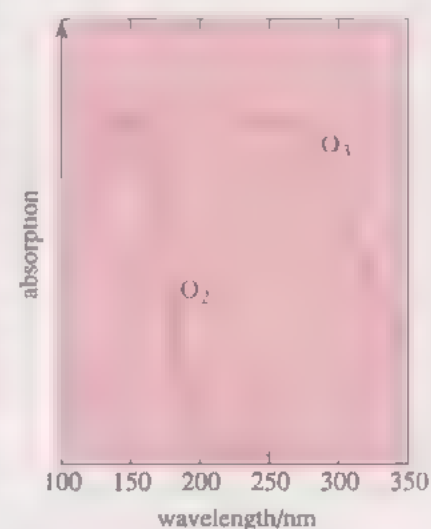


Figure 6.9 The absorption spectra of oxygen (O_2) and ozone (O_3) in the ultraviolet region of the electromagnetic spectrum.

The phases of matter are discussed in Section 2.3 of *Preparatory Science*.

The basis of the prediction of the phase in which a substance will exist is the **saturation vapour pressure diagram** (see Box 6.1), which can be constructed from laboratory studies of the substance in question.

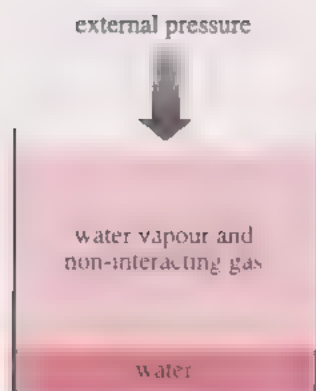


Figure 6.10 A model of an atmosphere above a pool of liquid water. In this model it is possible to vary the pressure and the temperature, factors that determine the amount of water vapour in the atmosphere.

It is the number density of *particles* that is relevant here, as in Book 1, Section 3.2.

Box 6.1 Saturation vapour pressure diagrams

As an example of a partial pressure diagram we shall consider how rain forms in an atmosphere that contains water vapour. Because the atmosphere of the Earth or any other planet is complicated by the presence of various zones and by atmospheric motion, we shall develop the diagram for a much simplified model system. Our model consists of a closed box that contains water and a gas (for example dry air) that does not interact with the water (Figure 6.10). Because the box is closed, nothing leaves or enters it. However, the lid of the box can move up and down to change the pressure in the box. This movement simulates the effect of changing the total pressure in the Earth's atmosphere. We can also raise and lower the temperature.

Within the box, water is present as a liquid and also as a vapour, these two phases being at equilibrium, a situation represented by Equation 6.3:



in which (l) represents the liquid phase and (g) represents the gaseous phase. At equilibrium, there is no change and so the ratio of the number densities, $n(\text{H}_2\text{O}, \text{g})/n(\text{H}_2\text{O}, \text{l})$, of the two phases is constant. As you might expect, this ratio depends on the temperature and the pressure, and it is this dependence that concerns us.

As a liquid, water is not easily compressed, and so the number density $n(\text{H}_2\text{O}, \text{l})$ remains nearly constant if the pressure is changed. The number density of water in the gaseous phase, $n(\text{H}_2\text{O}, \text{g})$, does change with pressure because gases are relatively easily compressed. So at a particular temperature, we need consider only how $n(\text{H}_2\text{O}, \text{g})$ changes with pressure. This number density is proportional to the *partial* pressure exerted by those molecules. It is therefore convenient to represent the equilibrium ratio in Equation 6.3 by the partial pressure of $\text{H}_2\text{O}(\text{g})$ that is in equilibrium with $\text{H}_2\text{O}(\text{l})$. A plot of the variation of the equilibrium partial pressure with temperature is called a saturation vapour pressure diagram, and that for H_2O is shown in Figure 6.11. The upper part of the curve (above the point T) represents $\text{H}_2\text{O}(\text{g})$, in equilibrium with $\text{H}_2\text{O}(\text{l})$. Below the point T the curve represents $\text{H}_2\text{O}(\text{g})$ in equilibrium with solid water, ice, $\text{H}_2\text{O}(\text{s})$.

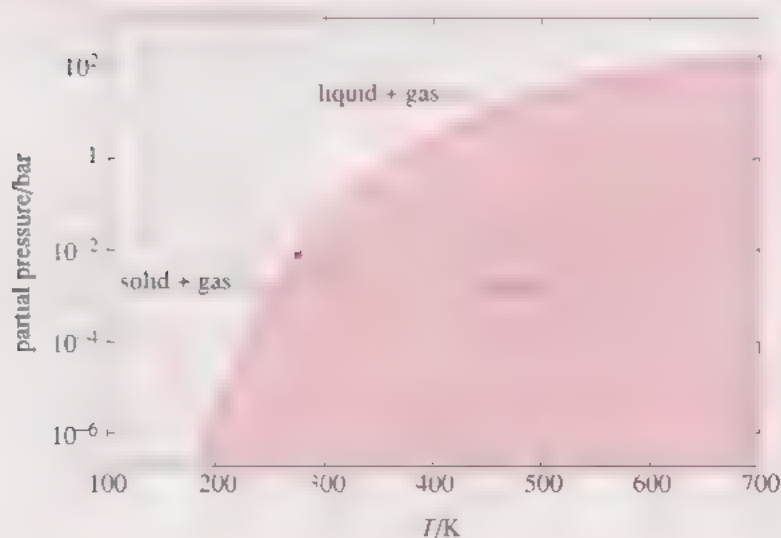


Figure 6.11 The saturation vapour pressure diagram for water. The pressure scale is logarithmic. Note that 1 bar = 10^5 Pa, close to atmospheric pressure at the Earth's surface.

So the curve in Figure 6.11 gives the value of the partial pressure of water vapour that is in equilibrium with liquid water or ice in the range of temperature shown. If the box in Figure 6.10 and its contents are cooled, the partial pressure of water in the gas at equilibrium with liquid water will be lowered, as Figure 6.11 shows. At this lower temperature, the gas will therefore contain initially *more* water than the equilibrium concentration, i.e. the partial pressure of water in the gas will exceed that at equilibrium. The result is that some of this water vapour will condense to liquid until the partial pressure of water vapour in the gas reaches the equilibrium value, the value on the saturation vapour pressure curve at the lower temperature.

If the equilibrium partial pressure and temperature are to the *right* of the saturation vapour pressure curve, then the substance is present as a gas only (Figure 6.11). If the pressure or temperature is then changed to lie on this curve, some liquid or solid will form. This condensation of water vapour to liquid or solid in the gas is the same process by which clouds form in the Earth's atmosphere.

From Figure 6.11 we can predict whether water vapour will condense in the atmosphere.

Suppose that on a warm, humid day the partial pressure of $\text{H}_2\text{O}(\text{g})$ at ground level rises to 10^{-2} bar, and then the temperature falls abruptly at night to -10°C (263 K). Will water vapour condense in the atmosphere and, if so, in what phase?

■ As the temperature falls, the point representing the gas moves horizontally across the 'gas only' region of Figure 6.11, until it meets the curve. At a pressure of 10^{-2} bar, this is at a temperature of about 275 K. Thus, condensation will begin at this temperature, and the partial pressure will fall until the point on the curve at 263 K is reached. The point on the curve corresponding to 263 K is below the point T, so water will condense as ice. [Comment: These are the conditions under which freezing fog forms at ground level on Earth.]

A cloud is a region of the atmosphere in which a component has condensed to form small liquid droplets or solid particles. These droplets or particles are so small that they are held suspended by updraughts in the cloud, until they grow large enough to fall as rain or snow. Saturation vapour pressure diagrams, such as Figure 6.11, enable us to predict the equilibrium conditions when clouds will form, but such predictions give only an approximate estimate as they ignore some important effects.

The most important factor that affects cloud formation is the shape of the droplets. Figure 6.11 applies to a flat liquid surface, as in Figure 6.10. At any particular temperature, the partial pressure of vapour in equilibrium with the curved surfaces of small spherical droplets is higher than that for a flat water surface, and it increases as the drop size decreases (as curvature increases). So at the initial formation of very small droplets, a *higher* partial pressure is needed than that predicted by saturation vapour pressure diagrams. However, atmospheres are not simply gaseous mixtures. They contain particles of dust, on which water molecules may condense to begin the process of drop formation. If the surface of these particles is easily wetted, and especially if the particle is soluble in water, the drop will form and grow even at partial vapour pressures *lower* than those predicted by saturation vapour pressure diagrams. Like many aspects of atmospheric science, cloud formation is not highly predictable.

The clouds of Venus (Plate 2.29) cover the entire planet at all times between 45 and 65 km altitude. They are composed largely of droplets of

concentrated sulphuric acid (H_2SO_4), whose chemical formation is considered in Section 6.5. The high surface temperature of the planet ensures that sulphur is vaporized on Venus and forms the source of the H_2SO_4 . The ability of the top surface of the cloud layer to reflect solar radiation makes for a high albedo and keeps the effective temperature of Venus low.

Cloud formation, of liquid and solid H_2O , is highly variable on Earth. Typically about half of the planet is covered by cloud at any time.

On Mars, clouds consist of small crystals of ice (H_2O or CO_2). The ice clouds form locally because of the large diurnal temperature range (see Plate 2.31) and in winter as a thin sheet in the polar regions. Dust clouds raised by wind also occur, both locally and occasionally on a planet-wide scale.

6.4.4 Atmospheric motion

The troposphere is described as the region of atmospheric mixing, which includes horizontal motion of the atmosphere (wind). On Earth, this motion is variable and turbulent. In this subsection, we concentrate on the major processes that produce a general rather than a local circulation of atmospheres. On the global scale, we recognize several different kinds of major process for transport in the atmosphere. On terrestrial planets, the factors responsible are the heating of the planet by the Sun and the rotation of the planet.

On Earth, the most important atmospheric circulation arises from the differential heating by the Sun due to the spherical shape of the planet. Near the Equator, the Earth's surface receives more energy per unit area than it does near the poles (Figure 6.12). A beam of solar radiation is spread over a larger area near the poles than at the Equator. The higher temperature of the Earth's surface at the Equator causes warm air to rise vertically in the troposphere, cooling as it rises, and then to travel towards the poles. As it moves polewards, this air radiates heat and so cools further, until it eventually sinks back to the surface where it returns to the Equator, completing a cycle that is called a **Hadley cell**. These cells extend to about 30°N and 30°S , as shown in Figure 6.13.

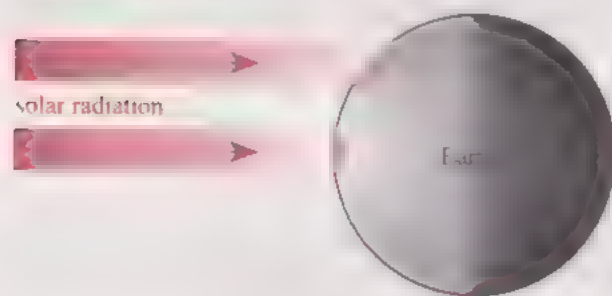


Figure 6.12 The Earth receives more solar energy per unit surface area near the Equator than it receives near the poles

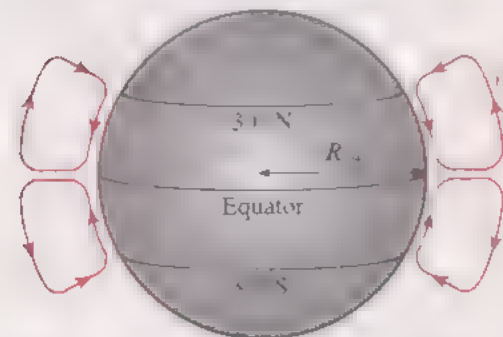


Figure 6.13 Hadley cells in the Earth's tropics extend to about 30°N and 30°S (not to scale).

The fast rotation of the Earth (rotation period of one day) has an additional effect on the atmospheric circulation through the interaction of the rotation with Hadley cell motion. We can understand this effect by following the motion of a piece of atmosphere around a Hadley cell as it circulates on the rotating Earth. Suppose that this piece of atmosphere begins on the surface of the Earth at the Equator, with no initial motion relative to the surface. It is rotating with the planet, with some west to east speed v relative to a non-rotating observer in space, moving at the same speed as the Earth in its orbit. This piece of atmosphere rises, without significant change in v (the height of the Hadley cell in Figure 6.13 is exaggerated for clarity). As it moves north, at the top of the cell, the piece of atmosphere, of mass m , does not change height, but its distance r from the Earth's spin axis decreases. At 30° N, its distance from the spin axis is now $R_\oplus \cos 30^\circ$ (Figure 6.14), where R_\oplus is the Earth's radius. As angular momentum (magnitude mvr) is conserved during this motion in which r decreases, it follows that v must increase (Book 1, Subsection 4.4.1).

At the Equator, the west to east speed of the atmosphere at the top of the cell, which we assume is equal to the speed of the surface, is $2\pi R_\oplus/T$, where T is the rotation period of the Earth. R_\oplus is 6378 km and T is 8.617×10^4 s (sidereal period). So this speed relative to the observer in space is 465 m s^{-1} .

11Q 6.4 What will be the west to east speed of the atmosphere at the top of the Hadley cell at 30° N relative to the observer?

Because the Earth is a solid body, at 30° N the angular speed of the surface is the same as at the Equator. At this latitude, the west to east surface speed is therefore $2\pi R_\oplus \cos 30^\circ/T = 403 \text{ m s}^{-1}$. The top of the Hadley cell thus acquires a west to east motion *relative to the surface*, as shown in Figure 6.15.

So we conclude that, at 30° N, the top of the Hadley cell has acquired a west to east speed relative to the surface of $(537 - 403) \text{ m s}^{-1} = 134 \text{ m s}^{-1}$. If, again, no significant change in v occurs as this piece of atmosphere returns to the surface at the northern extremity of the Hadley cell, the calculation predicts a wind speed west to east of 134 m s^{-1} , which is very high. The calculation ignores the effects of friction and small eddy currents, which reduce this wind speed, although the prevailing wind direction is accounted for. This wind is an example of the **Coriolis effect** – a displacement arising as matter moves in a rotating system.

Each of the two tropical Hadley cells forms a spiral of ascending and descending air that extends around the Earth (see Figure 6.16). At higher latitudes, the Hadley circulation becomes unstable, with the result that the tropical cells extend only to 30° N and S. At higher latitudes still, two further sets of less prominent and less permanent Hadley cells occur in each hemisphere (Figure 6.17), the intermediate one being driven by neighbouring cells *against* the direction suggested by the temperature difference that occurs with latitude.

The wind patterns established by Hadley cells have been known to travellers for centuries. At the Earth's surface the northern equatorial Hadley cell results in a general north-east to south-west motion of the atmosphere (Figure 6.17) giving the trade winds. Near the Equator, where the two tropical Hadley cells abut, an almost windless region exists, known as the doldrums. The top of the most northerly Hadley cell has a strong west to east motion known as the jet stream. At the high altitudes of over 10000 m at which transatlantic jet aircraft fly, the jet stream ensures that aircraft make the eastbound journey faster than the westbound.

The Earth rotates quite rapidly, causing a large Coriolis effect that limits the extent of Hadley cells. We can compare atmospheric motion on Earth with that on Venus and Mars. Venus rotates very slowly, once every 243 days, in a retrograde direction.

Eastward
West - east

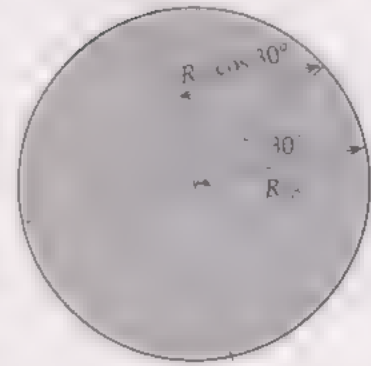


Figure 6.14 The air in a Hadley cell moves closer to the Earth's spin axis as it travels towards the pole. At 30° N the distance has decreased from R_\oplus to $R_\oplus \cos 30^\circ$.



Figure 6.15 As it moves north, the top layer of the Hadley cell acquires an easterly motion relative to the surface of the Earth. This is a manifestation of the Coriolis effect.

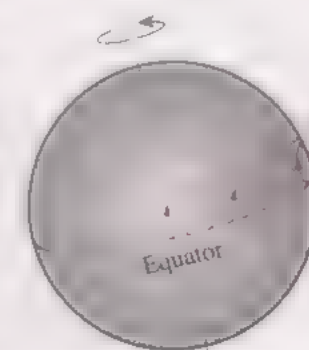


Figure 6.16 The Earth's rotation causes the Hadley cell to spiral. A piece of atmosphere that remains in the Hadley cell follows this flattened and tilted spiral path. This figure shows part of the tropical cell in the northern hemisphere; the vertical component is exaggerated.

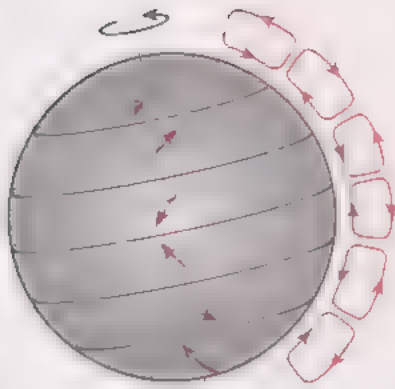


Figure 6.17 On Earth, three Hadley cells occur in each hemisphere. The tropical cells are the most persistent. The intermediate cells are driven by neighbouring cells in a direction contrary to the thermal gradient at the surface. The directions of surface winds are shown by arrows.

In which direction, west to east or east to west, would you predict the atmosphere at the top of the Hadley cells on Venus to move?

■ As the rotation of Venus is retrograde (east to west), this is also the predicted direction of the atmospheric motion.

How would you expect this slow rotation of Venus to affect the east to west movement of Hadley cell circulation?

■ The small Coriolis effect should contribute little to this movement, and so east to west motion might be expected to be slow (cf Figure 6.15).

At visible wavelengths, the cloud tops of Venus are almost featureless, but imaged in UV radiation (Plate 2.29) they show cloud patterns that indicate that Hadley cells extend from the equator to the poles. Heat is thus transported from the tropics to the poles, resulting in a relatively small surface temperature difference (about 10 K) between equator and poles, compared with 45 K on Earth. As might be expected on a slowly rotating planet, wind speeds at the surface are low, only about 1 m s^{-1} , as measured by entry probes and landers. However, at the cloud tops (altitude 65 km) the wind speed is typically 120 m s^{-1} in the retrograde direction of the rotation of the planet. This phenomenon is called **super-rotation** and its cause is not well understood, but is probably related to the fact that most of the solar radiation is absorbed at high altitude in the atmosphere of Venus. This is in contrast to what happens on Earth and Mars, where most of the solar radiation reaches the surface.

The rotation period of Mars is similar to that of the Earth, so it is not surprising to find a similar pattern of atmospheric circulation. Hadley cells occur near the equator but, owing to the thin atmosphere, the transport of heat from the tropics to the poles is less effective. Surface temperatures of more than 240 K were recorded by the Viking 1 lander at 22° N , with a diurnal variation of up to 50 K. The temperature at the poles drops to less than 150 K, sufficient for CO_2 , the major atmospheric component, to solidify. In Plate 2.32 this condensation is evident in the white polar cap, which consists of solid CO_2 and H_2O .

The condensation of CO_2 on Mars, together with the inclination of its spin axis, leads to a further type of atmospheric circulation. Consider the inclination first. The angle of inclination of the spin axis is similar to that of the Earth, about 24° . Seasons occur, each hemisphere alternating between summer and winter (see Box 6.2).

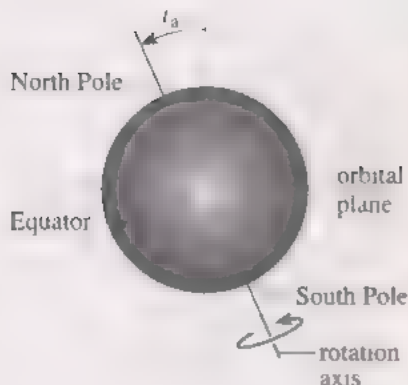


Figure 6.18 The axial inclination i_a of a planet.

Box 6.2 Seasons

On Earth, we are familiar with alternating, annual **seasons** of summer and winter. Seasons are caused by the angle that the spin axis makes with the normal to the orbital plane. This angle is called the axial inclination (Figure 6.18); for both Earth and Mars it is about 24° .

In the northern hemisphere summer, that hemisphere is tilted towards the Sun (Figure 6.19) with the result that day is longer than night and the Sun is high in the sky. At this time it is winter in the southern hemisphere and the opposite conditions prevail. When the Earth has completed a further half orbit about the Sun, the seasons are reversed.

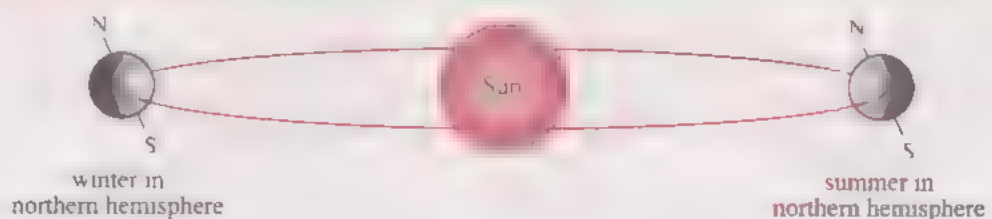


Figure 6.19 The seasons alternate as the Earth orbits the Sun (not to scale).

With the arrival of summer in the Martian northern hemisphere, the rise in temperature causes the polar CO_2 ice-cap to evaporate. In fact, CO_2 passes directly from solid to gas by sublimation. At the same time, the temperature at the south pole decreases and CO_2 solidifies from the atmosphere, increasing the size of the polar cap. This phenomenon causes a flow of atmosphere from one pole to the other, a flow that is reversed with the Martian seasons, and is called the **condensation flow** (Figure 6.20).

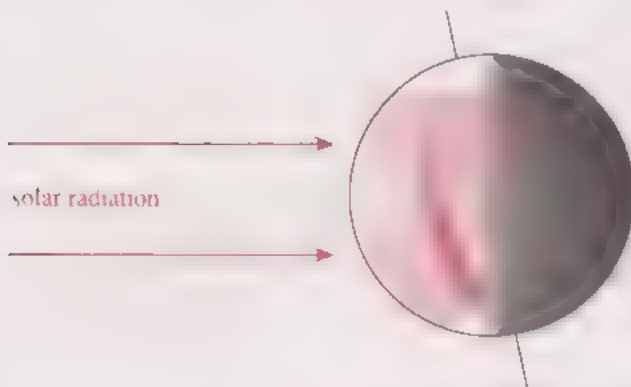


Figure 6.20 During summer in the Martian northern hemisphere, condensed substances (mostly CO_2 with some H_2O) evaporate and migrate to the southern pole, where they condense as ices.

A further type of atmospheric motion that has been detected on Mars can be attributed to the thinness of the Martian atmosphere. During daytime, the surface heats up rapidly because very little solar radiation is absorbed by the overlying atmosphere. At night, radiation from the surface readily escapes from the planet, again because the atmosphere is so thin. The result is a large difference in temperature between day (220 K) and night (170 K). This temperature difference causes a flow of atmosphere around the planet, referred to as a **thermal tide**, although it is not really a tidal process (see Figure 6.21).

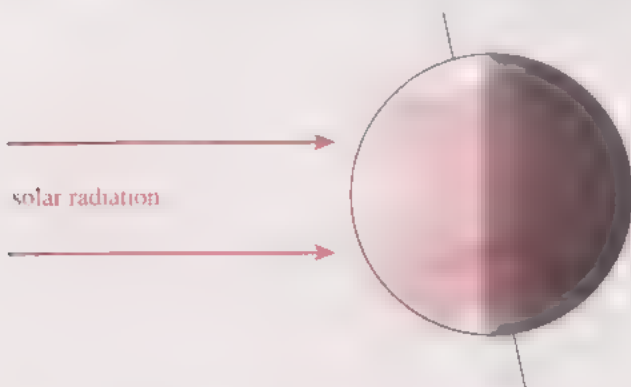


Figure 6.21 On Mars, a diurnal motion is caused by the large difference in temperature between the day and night sides of the planet.

The combination of these various atmospheric motions and local eddy currents on Mars can lead to local surface wind speeds as high as 50 m s^{-1} . These winds raise dust from the surface, creating local dust storms. Within these storms the dust absorbs solar radiation, setting up large temperature differences and further turbulence by a process of positive feedback. Over a Martian year many local dust storms occur on Mars, and once or twice a year the whole planet becomes enveloped in a storm.

It would now be useful to view video sequence 7 Atmospheric circulation (but leave the exercise until another viewing in Chapter 7). Remember to read the associated notes first.

Summary of Section 6.4 and SAQs

- 1 The balance between incoming solar radiation and the thermal emission by a terrestrial planet determines the average temperature of its atmosphere and surface. The greenhouse effect causes average surface temperatures to be higher than effective temperatures.
- 2 The atmospheres of Venus, Earth and Mars show vertical variation of temperature, density and pressure. On this basis, the following regions of the atmosphere can be distinguished:
The troposphere is the lowest region, in which surface heating causes convection to occur;
The stratosphere, unique to Earth, is the middle region, heated by the absorption of solar UV radiation by ozone, O_3 ;
The thermosphere is the highest region, and is heated by direct absorption of solar UV radiation, mainly by O_2 (Earth) and CO_2 (Venus and Mars).
- 3 Cloud formation occurs on Venus mainly as H_2SO_4 droplets, on Earth as water droplets and ice particles (H_2O), and on Mars as ice (H_2O and CO_2) and dust particles.
- 4 Motion in the troposphere circulates the atmosphere. Hadley cells with associated west to east Coriolis motion occur on Earth and Mars. Venus rotates slowly, yet its single Hadley cell structure shows high east to west wind speeds at high altitudes (super-rotation). On Mars, the atmosphere further circulates by condensation flow (pole to pole) and by a thermal tide (flow from day-side to night-side).

SAQ 6.3 (Objectives 6.1 and 6.3) Triton, the largest satellite of Neptune, has a high albedo. What does this indicate about its surface? Given that it is a long way from the Sun, is its low surface temperature of 38 K surprising?

SAQ 6.4 (Objective 6.4) The release of chlorofluorocarbons (CFCs), which were once commonly used as aerosol propellants, into the Earth's atmosphere introduces substances that absorb infrared radiation in a different part of the spectrum from that absorbed by CO_2 . What effect would you predict this to have on the temperature of the Earth's atmosphere?

SAQ 6.5 (Objective 6.5) Using Figure 6.22, determine whether CO_2 would condense as a solid or a liquid on Mars at its average surface temperature.

SAQ 6.6 (Objectives 6.1 and 6.6) What speed relative to the surface will a piece of atmosphere gain by the Coriolis effect in travelling on Venus from the equator to $30^\circ N$, and in which direction at the top of the Hadley cell will that gain be? (Take the surface speed of the rotating planet at its equator to be 0.997/243 of that of the Earth, as the two planets are of similar size and Venus rotates once every 243 Earth (solar) days in the retrograde direction, east to west.)

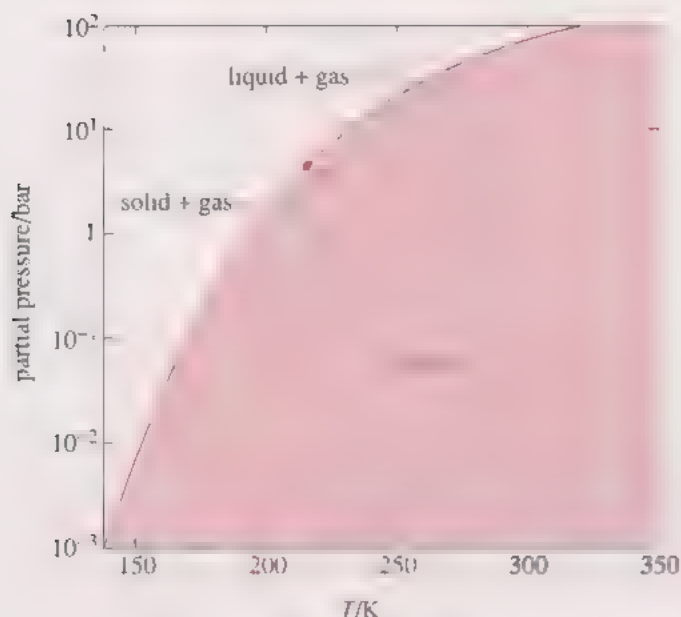


Figure 6.22 The saturation vapour pressure curve for carbon dioxide, CO_2 .

6.5 Atmospheric chemistry

6.5.1 Chemical reactions

At present, the gross composition of the Earth's atmosphere does not appear to be changing significantly. There is no evidence to suggest that the atmospheres of Mars and Venus are changing either. However, this is not to suggest that these atmospheres undergo no long-term change. It is known that the atmosphere of the Earth, as deduced from the geological record, has undergone major change (see Section 6.6). It is very likely that the atmospheres of Venus and Mars have also changed since the formation of the planets. In spite of this apparent short-term stability, the Earth's atmosphere is far from being an equilibrium mixture, that is, a mixture of the composition that would be attained if the constituents were isolated in a closed box for a *very* long time. A measure of the departure from chemical equilibrium is the co-existence of the two major components, nitrogen and oxygen. In the presence of the current partial pressure of oxygen, the *equilibrium* partial pressure of nitrogen (N_2) is less than 10^{-10} bar, whereas the *actual* partial pressure of N_2 is about 0.8 bar.

The apparent stability of the atmospheres is thought to be the result of a **steady state**, in which the rates of the various reactions that produce each component are balanced by the rates of the reactions that remove it, without any of these reactions ever reaching equilibrium.

If we wish to understand the chemistry of these atmospheres, it is necessary to know which reactions are important. On the basis of a knowledge of individual reactions, we can construct a model of the major processes that take place. In the same way, our understanding of the nucleosynthetic processes that occur in stars is based on a knowledge of individual nuclear reactions (Book 1, Chapter 3). As with nuclear reactions, the individual chemical reactions, including the photochemical changes that are initiated by solar radiation, are studied in the laboratory to determine their rates and to deduce which are important. Before we can consider the chemistry of the three planets, some background on chemical reactions is necessary.

Box 6.3 Chemical reactions

Some general aspects of chemical reactions

In **chemical reactions**, atoms, ions and molecules are rearranged to produce different atoms, ions and molecules. Elements are neither created nor destroyed in a chemical reaction, nor is there any change in the isotopes. Thus nuclides are conserved, and it is this conservation that distinguishes chemical from nuclear reactions. On the other hand, as in a nuclear reaction, a chemical reaction must be balanced. In this case, each of the atoms present in the reacting substances (the reactants) must appear in products of the reaction. For example, consider the reaction of an atom of oxygen with a hydrogen molecule, which is shown pictorially in Figure 6.23: an atom of oxygen collides with a molecule of hydrogen to produce the hydroxyl molecule (OH) and an atom of hydrogen.

Chemical reactions are more usually represented by balanced chemical equations, in this case



Figure 6.23 A schematic depiction of the reaction of an atom of oxygen with a molecule of hydrogen.

Write a balanced chemical equation to represent the reaction of an oxygen atom with a water molecule (H_2O) to produce the hydroxyl molecule

In addition to a change in the chemical identity of the reacting substances, **energy changes also occur in chemical reactions**. Reaction 6.4 occurs in atmospheres, in the gas phase, and so the initial requirement is for an O atom and a H_2 molecule to collide. However, collisions of this kind do not always yield products. If the relative kinetic energy of the colliding O and H_2 is low, they simply bounce off each other. Only if their relative kinetic energy is greater than some critical value does reaction ensue. This requirement is interpreted as an **energy barrier** that must be surmounted for reaction to occur. If the kinetic energy is less than the energy barrier, the colliding partners just bounce apart. If it is greater than the energy barrier the atoms intermingle to produce the new substances, the products OH and H.

For Reaction 6.4, the potential energy in the chemical bond O—H is greater than that in the chemical bond H—H. This potential energy is called **chemical energy**. Notice that there is no chemical energy in an isolated atom. So as a mixture of O and H_2 reacts, it draws on the kinetic energy of the reactant molecules, converting this energy into potential energy. The result is that the system cools. To keep the reaction going, energy must be supplied to the reacting gases. Such a reaction is called an **endothermic reaction**, and all such reactions have an energy barrier.

The reverse reaction can also take place:



Obviously this reverse reaction raises the kinetic energy (and temperature) of the gaseous mixture as it proceeds. It is called an **exothermic** reaction. Although the *net* effect of an exothermic reaction is an increase of kinetic energy, there may still be an **energy barrier to overcome**. In fact, most chemical reactions have an energy barrier.

The **rate of reaction** is the number of occasions that the reaction occurs between molecules in unit time (per second). This rate increases as the rate of collisions increases. The **collision rate** depends on two factors. It increases as the number density of each of the reactants increases. It also increases with the speed of the reactant molecules. An increase in the speed of the reactant molecules, i.e. in their kinetic energy, also assists in overcoming the energy barrier. So, because kinetic energy increases with temperature, reactions generally proceed faster as the temperature is raised.

If we start with a certain amount of reactants, O and H₂, and keep the temperature of this mixture constant, the products will build up until the rate of Reaction 6.5 equals the rate of Reaction 6.4. We may then write



This represents **chemical equilibrium**, in that the amount of each reactant and product no longer changes. This does not mean that Reactions 6.4 and 6.5 have stopped. Equation 6.6 represents a dynamic process, in which the rates of the forward and reverse reactions are equal so that no net change occurs.

Reaction 6.4 involves the transfer of one H atom from one atom (H) to another (O) with the breaking of one bond and the formation of another. In some reactions only bond formation occurs, in others bonds are only broken. Here we consider different kinds of chemical reaction.

Reactions in which molecules dissociate

When a molecule breaks into smaller fragments, it is said to dissociate. In the **dissociation** of a water molecule, for example, the chemical bond that holds a hydrogen atom and an oxygen atom together is broken. Chemical bonds generally consist of two electrons, the **electron pair bond**. In this kind of dissociation, the chemical bond is broken in such a way that one of these electrons ends up on hydrogen and the other on oxygen. The result is a hydrogen atom and a hydroxyl molecule. Hydroxyl is a special type of molecule that contains an *odd* number of electrons, the odd electron being called an unpaired electron. Such molecules are called **radicals**. Again, this reaction can be represented by a sketch, as in Figure 6.24. It can also be represented by an equation:



Figure 6.24 A schematic depiction of the dissociation of a water molecule, H₂O.

Note the difference between this dissociation and that of H₂O in the liquid phase. Here the electron pair is split between H and OH, two neutral fragments. In liquid H₂O the electron pair ends up on the oxygen, leaving the hydrogen with a positive charge as in H⁺, the hydrogen ion, and the oxygen with a negative charge as in OH⁻, the hydroxide ion. The dissociation of H₂O into H and OH (the radicals) is called **homolytic dissociation** or **homolysis**. Sometimes the presence of the odd electrons that result from homolytic dissociation is represented by a dot, i.e. the products of the reaction are written as H[•] and OH[•].

Reaction 6.7 is an endothermic reaction, as are all dissociations in the gas phase. In general, molecules, in this case the water molecule H_2O , are stable. They do not simply fall apart. Energy is needed to break bonds. In planetary atmospheres, this energy can be provided by one of two means. It can be thermal energy due to the temperature of gas. Thermal dissociation is much more important in the outer planets than in the terrestrial planets. It can also be provided by the energy of a photon. A photon of certain energy, when absorbed by a molecule of H_2O , causes dissociation, the photon providing the energy needed to break the bond. We represent this by:



A reaction in which a molecule dissociates as a consequence of absorbing a photon of electromagnetic radiation is called **photodissociation** (or sometimes photolysis). The general term for the overall result of such a reaction and the ensuing reactions of the radical products is **photochemistry**.

The amount of energy needed to break a chemical bond depends on the nature of the bond being broken. It lies roughly in the range from 5×10^{-19} to 2×10^{-18} J per bond.

In which part of the electromagnetic spectrum do such photon energies occur? (Energy is related to frequency by $E = hf$, see Book 1, Subsection 1.2.4.)

From the relationship $E = hf$, we deduce that these energies correspond to photon frequencies in the range from 8×10^{14} to 3×10^{15} Hz, the UV region.

So solar UV radiation is capable of causing the photodissociation of all molecules. Indeed, this happens in the upper parts of all atmospheres, and often the absorption of UV radiation is complete, thereby protecting the lower atmosphere from photodissociation.

Reactions in which bonds form

When two atoms collide, you might suppose that they would combine to form a molecule that is stable, for example in the encounter of two oxygen atoms to form a molecule O_2 :



This **bond formation** is the reverse of the dissociation of O_2 , in which energy is required to break the bond. Here the reaction is exothermic, that is energy must be released if the bond is to form, and so there must be some way of disposing of this energy. For small molecules, the energy cannot be accommodated internally by spreading it amongst the vibrations of the molecule as this will only lead again to dissociation. This is not a problem with large molecules where there are many vibrations amongst which the energy may be shared. For small molecules, however, it is necessary to dispose of the energy to a so-called **third body**, another atom or molecule involved simultaneously in the collision, for example a nitrogen molecule, N_2 , that does not react with oxygen atoms. This third body is usually represented by a general symbol M. The combination of two oxygen atoms is then written as



Of course, these **three-body collisions** occur less frequently than collisions between two atoms, with the result that this type of reaction is often slow.

Reactions involving bond breaking and bond making

Chemical reactions often involve both the making and breaking of chemical bonds, as in the reaction of O with OH, which is involved in the overall photodissociation of H₂O to produce oxygen and hydrogen:

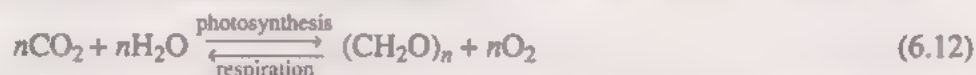


Some of these reactions are endothermic, as is manifest by their tendency to proceed faster at higher temperatures. Energy is required if, as in this case, the energy of the bond broken is greater than the energy of the bond formed. Even where the reverse is true, these reactions tend to be faster at higher temperatures. This is because the process involves partial breaking of one bond (requiring energy) before there is compensation from the bond that is to be formed.

6.5.2 Chemistry of the Earth's troposphere

The detailed composition of the Earth's troposphere is given in Table 6.5, which also lists the main sources and **sinks** for each component. The term sink describes the process by which a substance is removed from the atmosphere. Notice that biological control is the major source and sink for the two main components, oxygen and nitrogen.

The Earth's atmosphere is unique in the Solar System in containing a substantial amount of oxygen. The origin of oxygen is in **photosynthesis** by plant life, and its consumption occurs largely by plant respiration (Equation 6.12) and the decay of plant tissue after death. In this way, oxygen is linked with the balance of carbon-containing compounds.



Here, the formula (CH₂O)_n represents the range of carbohydrate materials that make up plants.

The forward, photosynthetic step in Equation 6.12 uses the energy of photons (light) to convert CO₂ into what is called a reduced form of carbon (see Box 6.4 below) by making hydrogen from water available, and evolving molecular oxygen. This process depends initially on the absorption of light by plant pigments, especially the pigment chlorophyll. Figure 6.25 shows the absorption spectrum of chlorophyll. The peaks represent different electronic transitions in this large molecule. The absorption of the blue and red ends of the visible part of the spectrum gives the pigment its green appearance.

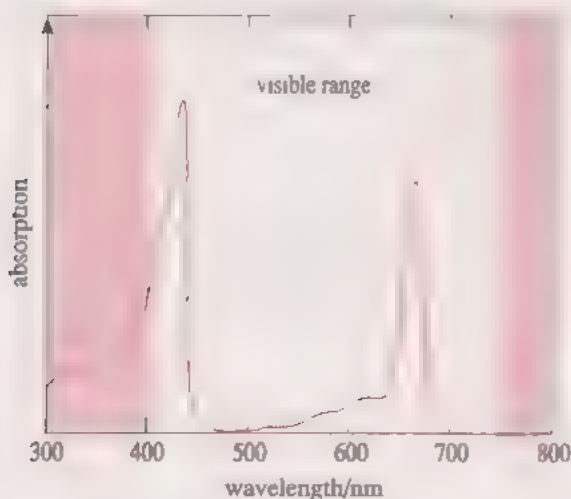


Figure 6.25 The absorption spectrum of the green plant pigment chlorophyll.

Table 6.5 The composition of the Earth's troposphere

Gas	Volume ratio ^a	Major source	Major sink
N ₂	7.81×10^{-1}	biology	biology
O ₂	2.09×10^{-1}	biology	biology
H ₂ O	$< 4 \times 10^{-2}$	evaporation	condensation
⁴⁰ Ar	9.3×10^{-3}	outgassing ^b	
CO ₂	3.4×10^{-4}	combustion, biology	biology
³⁶ Ar, ³⁸ Ar	3.7×10^{-5}	outgassing	
Ne	1.82×10^{-5}	outgassing	
He	5.24×10^{-6}	outgassing	escape
CH ₄	1.73×10^{-6}	biology	photo-oxidation
Kr	1.14×10^{-6}	outgassing	
H ₂	5×10^{-7}	photochemistry	escape (as H)
N ₂ O	3.1×10^{-7}	biology	photodissociation
Xe	8.7×10^{-8}	outgassing	
CO	4.2×10^{-8}	photochemistry	photochemistry
O ₃	1.1×10^{-8}	photochemistry	photochemistry
hydrocarbons	$\sim 1 \times 10^{-8}$	incomplete combustion	photo-oxidation
HCl	$\approx 1 \times 10^{-9}$	acidification	rainout
COS	5×10^{-10}	biology	photodissociation
CH ₃ Cl	5×10^{-10}	biology	photo-oxidation
SO ₂	$\approx 3 \times 10^{-10}$	combustion, photochemistry	photo-oxidation
CFCI ₃	2.2×10^{-10}	industry	photo-oxidation
H ₂ S	$\approx 2 \times 10^{-10}$	biology	photo-oxidation
NH ₃	1.1×10^{-10}	biology	photo-oxidation
HNO ₃	5.1×10^{-11}	photochemistry	rainout
NO ₂	2.5×10^{-11}	combustion, biology	photo-oxidation

^a Volume ratio is that fraction, by number of atoms and molecules, of the troposphere that is made up by the component listed.

^b Outgassing is discussed in Subsection 6.6.3.

Why is chlorophyll well suited to its role as a means of gathering energy from the Sun's radiation?

■ Comparison with Figure 6.7 shows that this absorption lies close to the energy maximum of the solar spectrum.

The photosynthetic process has been the subject of much investigation, as to reproduce it non-biogenically could provide a non-polluting and inexhaustible supply of energy by capturing CO₂ from the atmosphere and returning it by combustion.

Box 6.4 Oxidation and reduction

Oxygen occurs in combination with other elements in substances, such as H_2O and CO_2 , that are present in the Earth's atmosphere. In carbon dioxide, the ratio of oxygen to carbon (O : C) reaches its highest in compounds of these two elements. So in CO_2 , the carbon is described as an **oxidized substance**. When combined with the maximum amount of oxygen in a chemical compound, an element is described as oxidized.

The Earth's atmosphere also contains compounds of carbon in which there is no oxygen. In one example, methane, CH_4 , the ratio of hydrogen to carbon (H : C) has its highest value for compounds of these two elements: the carbon is now a **reduced substance**. When combined with the maximum amount of hydrogen, an element is said to be reduced.

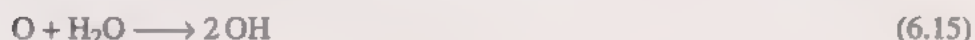
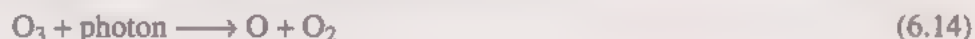
Oxygen, which is responsible for the conversion of carbon-containing substances into CO_2 , is an **oxidizing substance**. Hydrogen, which can convert carbon-containing substances to CH_4 , is called a **reducing substance**. The conversion processes are referred to, respectively, as **oxidation** and **reduction**. Intermediate levels of oxidation (or reduction) are represented by substances such as carbon monoxide, CO , which also occurs in the Earth's atmosphere. Oxidized and reduced are relative terms. For example, CO is oxidized relative to CH_4 , but reduced relative to CO_2 .

Present plant life is responsible for the major cycling of CO_2 , about 1.5×10^{14} kg of carbon annually. The cycle represented by Equation 6.12 is not closed. After death, most plant life is oxidized by decay processes, but some is buried and so its carbon is removed for long periods from the atmosphere. Atmospheric CO_2 dissolves in the oceans, but is also released by the oceans, the process being governed mainly by the prevailing temperature. This role of the oceans is the main factor in establishing a steady-state concentration of atmospheric CO_2 . Interactions with the Earth's land surface also remove CO_2 from and add it to the atmosphere by weathering. Buried organic material is brought to the surface and oxidized by weathering. The weathering of limestone (CaCO_3), a rock that is produced from the hard parts of microscopic plants and animals such as coral, also converts CO_2 into the soluble ion bicarbonate (HCO_3^-), which ends up in the oceans:

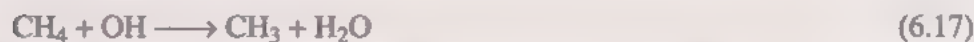


As Table 6.5 shows, several other carbon-containing molecules, especially methane (CH_4) and carbon monoxide (CO), occur in the troposphere, and are mostly of biological origin. The existence of these reduced forms of carbon in the presence of a high partial pressure of O_2 implies that the atmosphere is far from chemical equilibrium. These reduced forms of carbon are eventually oxidized, otherwise their concentrations would steadily increase. Their oxidation takes place mainly by oxygen atoms and by the hydroxyl radical, OH , which is formed from H_2O , which is always present.

Atomic oxygen is produced by the photodissociation of ozone, which occurs mainly in the stratosphere. However, ozone of stratospheric origin can reach the troposphere and be photodissociated there, to result in the formation of hydroxyl radicals by Equations 6.14 and 6.15:



In the troposphere, the dominant reactions are the oxidation of molecules of biological origin by OH. Of these, the most important are CO and CH₄:



Two of the products of these reactions, CO₂ and H₂O, are significant atmospheric components. The other two, H and CH₃ radicals, are highly reactive and combine rapidly with O₂. Further reactions produce substances such as hydrogen peroxide (H₂O₂) and CO₂, which dissolve in rain-drops. Reactions such as those given in Equations 6.16 and 6.17 are major processes that occur in the Earth's atmosphere; reduced or partially oxidized forms of carbon, generated biologically, are oxidized by OH to form CO₂.

The processes involving the carbon compounds described in this Section form part of the balance of atmospheric carbon chemistry, including interaction with biology and rocks, that is known as the **carbon cycle**. Figure 6.26 summarizes these processes.

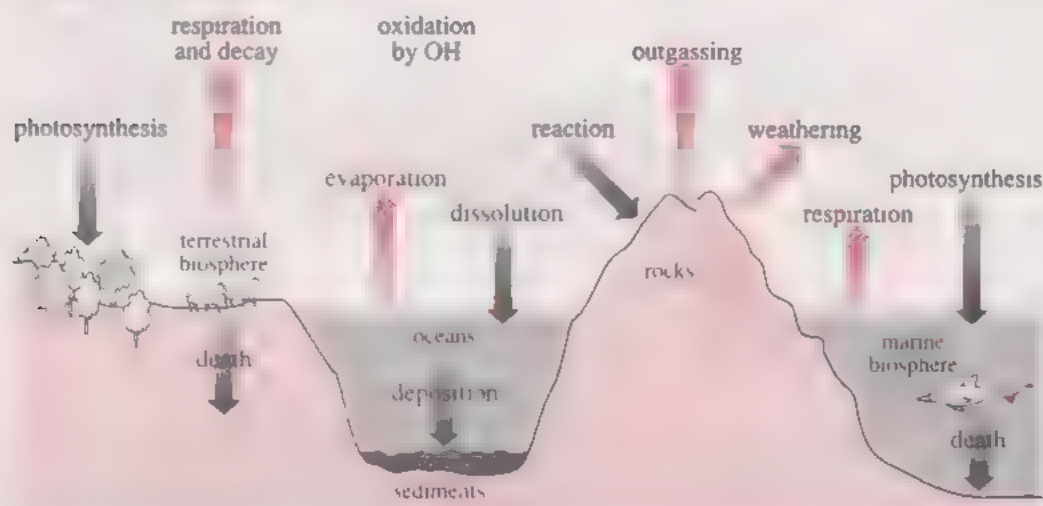


Figure 6.26 The carbon cycle on Earth.

We turn now briefly to the chemistry of nitrogen, another element that is cycled biologically in the Earth's atmosphere. The main nitrogen-containing molecule, N₂, also occurs on Venus and Mars; on Venus, it makes up about 2% by mass of the atmosphere.

How does the total N₂ compare on Venus and Earth?

■ The amounts of N₂ on the two planets are similar. On Venus, the atmospheric column mass is about 100 times that of the Earth (Table 6.3), of which about 2% is contributed by N₂. On Earth, most of the atmosphere (about 80%) is N₂. The surface areas of the two planets are similar.

N₂ is a rather inert molecule, as is N₂O, which is the second most abundant nitrogen-containing component of the Earth's atmosphere (Table 6.5). Other atmospheric compounds of nitrogen, such as HNO₃ and NO₂, are not highly abundant because they are much more reactive. They are produced by the reaction of ozone, O₃, with N₂O, and oxidize carbon compounds very effectively. HNO₃, which dissolves in water to produce nitric acid, is also removed from the atmosphere by rain.

Atmospheric nitrogen is cycled between those parts of the planet with which it has contact: the biosphere, the oceans and the soil.

The principal interaction with the biosphere is the conversion by bacteria of otherwise inert N₂ into NH₃ (ammonia) and thence the water-soluble ion NH₄⁺ (ammonium), a process called **nitrogen fixation**. Another plant nutrient, NO₃

Rain and rivers also deposit the ions NH_4^+ and NO_3^- in the oceans, where they are used by marine plant life. Again decay, and other processes, return nitrogen to the atmosphere. The Earth's lithosphere too is involved in the cycling of nitrogen. Buried biological matter removes nitrogen, which is eventually released, along with some 'new' nitrogen, as a result of geological activity.

11Q.66 In Figure 6.27, identify (i) the diatomic nitrogen compound X, formed as an intermediate in the conversion of N_2O into NO_2 , and (ii) the molecule Y, responsible for converting NO_2 into HNO_3 .



Within the stratosphere, ozone is produced from the interaction of solar radiation with molecular oxygen, O_2 . The chemistry of the stratosphere consists of a complex series of reactions involving O , O_2 and O_3 as well as a number of other atoms and molecules. One set of reactions involved in this chemistry is called the **Chapman scheme** after the English geophysicist Sydney Chapman (1888–1970), who first proposed it.

$$O + O_3 \longrightarrow 2O_2 \quad (6.21)$$

147

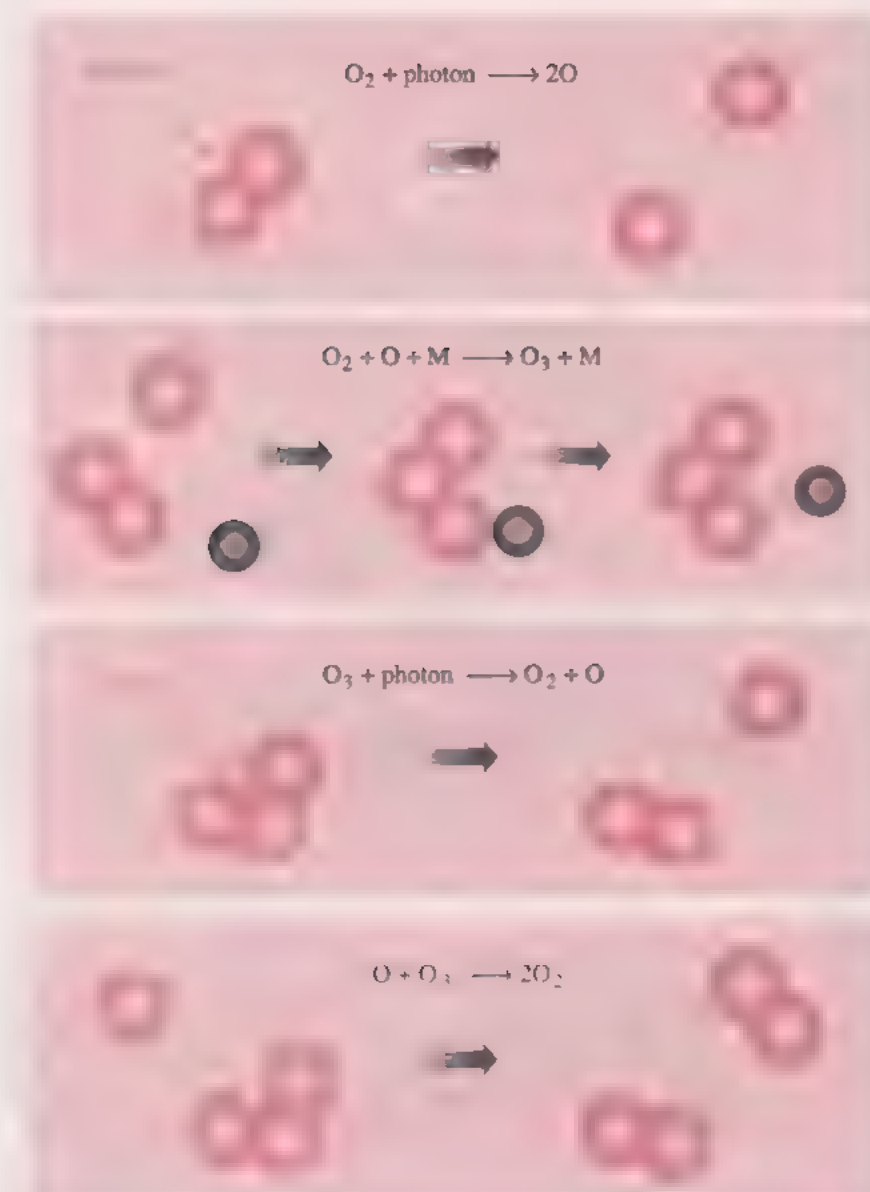


Figure 6.28 A schematic depiction of the four reactions of the Chapman scheme of ozone chemistry in the stratosphere. The body M (shaded grey) can be any atom or molecule in the three-body collision.

1. Q 6.7 What is the net effect of the Chapman scheme? You should deduce this by adding together all of the reactants and equating this with the sum of all the products. Cancel any atoms or molecules that occur on both sides of your equation. Notice that a molecule of ozone, O_3 , is a reactant in each of the Equations 6.20 and 6.21. As Equation 6.19 is the only reaction that produces ozone, you will need to count this reaction twice in the summation – this also takes care of the two oxygen atoms produced in Equation 6.18. Why is the general species M included on both sides of Equation 6.19?

The Chapman scheme explains satisfactorily the production of O_3 in the stratosphere, but accounts for only about 12% of its destruction. Ozone is destroyed by its reactions with other molecules and atoms, in particular the following: OH, oxides of nitrogen (NO_x as a general formula), and chlorine atoms (Cl). Much current interest focuses on the effects of pollution of the ozone layer through NO_x , generated by high-flying aircraft, and especially through Cl produced by the dissociation of chlorofluorocarbons (CFCs), which you met as a greenhouse gas in SAQ 6.4.

The Chapman scheme is a further example of several descriptions of atmospheric chemistry that we shall consider. As with all these schemes, this one concentrates on a small number of components. It attempts to explain the relationship between these components, how they are formed and how they are consumed in chemical reactions. Ozone exists at a particular concentration at any point in the atmosphere because it is formed and destroyed at some particular rate, leading to a steady-state concentration.

6.5.4 *Chemistry of the atmosphere of Venus*

The cloud layer of sulphuric acid on Venus is opaque (Plate 2.29), and divides the atmosphere into two distinct chemical regions. Above the clouds (> 65 km) the atmosphere is exposed to solar radiation, so chemistry at and above the clouds is dominated by the dissociation of molecules and by reactions of the reactive components that are produced. Within and below the clouds the atmosphere is shielded from solar radiation, although it is hot (up to 750 K at the surface). In this lower tropospheric region, reactions occur because of the relatively high temperature. We shall consider both regions, but first we examine how the cloud material H_2SO_4 , which is photodissociated by sunlight into SO_2 and H_2O , is regenerated.

The Venera landing craft detected numerous thunderstorms on their descent through the cloud layer of Venus. In an atmosphere that is substantially CO_2 with a few per cent of N_2 , lightning flashes produce oxides of nitrogen, NO and NO_2 . One possible route from SO_2 and H_2O to H_2SO_4 is via a sequence of reactions involving these oxides of nitrogen. Remarkably, the sequence of reactions involving NO , NO_2 , SO_2 and H_2O , which leads to H_2SO_4 , is thought to be the same as that in a nineteenth century process devised for the commercial production of sulphuric acid. In this process H_2O is consumed. In fact, sulphuric acid has a further dehydrating effect, and so H_2SO_4 droplets remove H_2O from the atmosphere at the cloud level, making it very dry.

The chemical composition of the troposphere of Venus is given in Table 6.6. At the high average temperature (730 K) of the surface, it might be expected that reactions would proceed rapidly towards an equilibrium mixture. However, the tropospheric chemistry is complicated by interactions with the surface and by convective transport of gas from the cloud layer. Notable amongst the compounds detected are several sulphur-containing substances. On Earth, many of the sulphur-containing molecules in the atmosphere are derived from biological sources, so it is surprising to find them in the atmosphere of Venus. Their origin is thought to depend on the high surface temperature and the occurrence of volcanism although, at the time of writing, volcanism has not been observed taking place on Venus (Subsection 5.6.4).

Volcanism would contribute sulphur vapour to the atmosphere from sulphide minerals, the sulphur being initially released as S_2 molecules under the conditions of temperature and pressure on Venus. Additionally, interaction of sulphur vapour and sulphide minerals with the main components of the atmosphere (CO_2 , CO and H_2O) leads to the formation of COS and H_2S (carbonyl sulphide and hydrogen sulphide) as well as SO_2 .

Within the atmosphere, SO_2 is more abundant than its equilibrium concentration, probably because of transport downwards from the cloud layer where SO_2 is produced, probably via SO_3 (sulphur trioxide), by photodissociation of H_2SO_4 .

Detailed chemical schemes have been proposed for many of the compounds in the lower atmosphere of Venus, including those containing Cl, N and C, but we have restricted our attention to the important sulphur compounds that are related to cloud formation and surface interaction. The sulphur cycle described here, together with further interconversions, is given in Figure 6.29.

Table 6.6 Composition of the troposphere of Venus

Gas	Volume ratio ^a	Major source	Major sink
CO ₂	9.65×10^{-1}	outgassing	CaCO ₃ formation?
N ₂	3.5×10^{-2}	outgassing	
SO ₂	1.5×10^{-4} (22 km)	photochemistry	CaSO ₄ formation
H ₂ O	1×10^{-4} (22 km) 1.4×10^{-6} (70 km)	outgassing, impacts	silicate hydration Fe oxidation plus H escape
COS	$< 4 \times 10^{-5}$	outgassing	photo-oxidation
^{36,38} Ar	3.7×10^{-5}	outgassing (primordial)	
⁴⁰ Ar	3.3×10^{-5}	outgassing	
H ₂	$< 2.5 \times 10^{-5}$	photochemistry	escape (as H)
CO	2×10^{-5} (22 km)	photochemistry	photo-oxidation
⁴ He	1.2×10^{-5}	outgassing	slow escape
Ne	7×10^{-6}	outgassing (primordial)	
H ₂ S	3.4×10^{-6}	outgassing	photo-oxidation
HCl	4×10^{-7}	outgassing	NaCl formation
Kr	5×10^{-8}	outgassing	
HF	5×10^{-9}	outgassing	CaF ₂ formation

^a Values at the surface unless indicated otherwise.

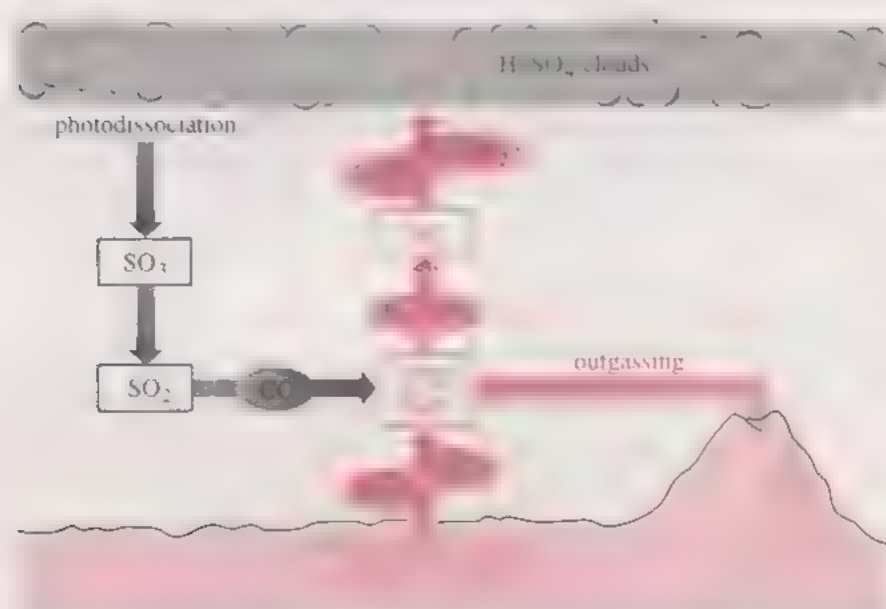
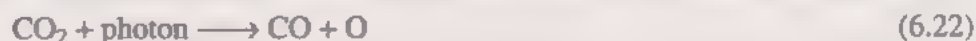
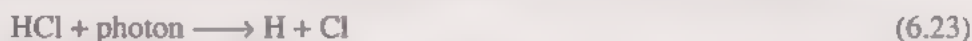


Figure 6.29 The sulphur cycle on Venus. SO₃ is known to be the product of the photodissociation of H₂SO₄, but its presence in the atmosphere is hypothetical.

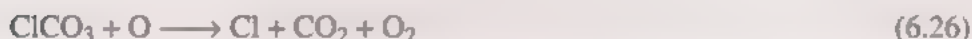
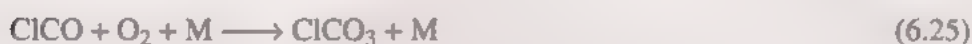
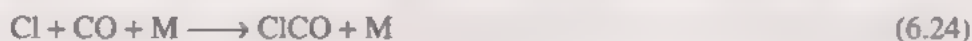
We turn now to examine the chemistry of the atmosphere *above* the clouds. Here, the atmosphere is exposed to solar UV radiation, which is capable of causing photodissociation of most of the compounds to be found in the upper atmosphere. Of these, CO₂ is the most abundant and dissociates to CO and atomic oxygen:



At these altitudes (> 65 km), however, CO is a very minor component. Nor is there evidence to suggest that carbon dioxide is being steadily depleted on Venus. So there must be a route to reverse the effect of Equation 6.22. The direct regeneration of CO₂ from CO and O is known to be a very slow reaction, but there are routes for this recombination that involve other chemical species. Amongst the molecules that have been detected in the upper atmosphere is hydrogen chloride, HCl. This too is dissociated by solar UV radiation:



Reaction of atomic chlorine with CO is the start of a sequence of reactions:



Notice that the first two steps are three-body reactions involving some unspecified molecule denoted by M.

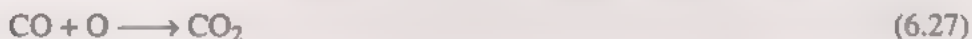
1 Why is the presence of M necessary in these two steps?

■ In both of these reactions there is a net formation of bonds, with the release of chemical energy. The need to dispose of this energy, which would otherwise cause the immediate disruption of the incipient product, is met by a third molecule (see Box 6.3).

On Venus, the probability of three-body collisions is relatively high owing to the high number density in the atmosphere, so the requirement of three-body collisions is not severe.

What is the net effect of Reactions 6.24 to 6.26?

■ Addition of left-hand and right-hand sides of the three reactions shows that the net effect is the recombination of CO with atomic oxygen to form CO₂:



This is just one of several routes that have been proposed for this recombination, as possible ways in which the CO₂ content of the atmosphere of Venus is maintained in spite of photodissociation of CO₂ by UV radiation.

11Q 6.8 In the Earth's troposphere, the hydroxyl radical, OH, derived from photodissociation of H₂O, plays a major role in oxidizing such molecules as CO to produce CO₂. Why does this process not play a significant role above the clouds on Venus?

6.5.5 Chemistry of the atmosphere of Mars

The atmosphere of Mars is composed mainly of carbon dioxide, with only a few per cent of N₂ and very minor amounts of other gases, including H₂O (Table 6.7). Being composed largely of CO₂, the atmosphere of Mars resembles that of Venus, especially at high altitude, although the column mass is very different. In addition to these atmospheric components, reservoirs of H₂O and CO₂ are contained in the polar ice-caps and permafrost, and also absorbed in the soil-like material on the Martian surface. So, not surprisingly, the Viking missions observed relatively large amounts of H₂O in the atmosphere close to the north polar cap, especially during summer when the cap is evaporating. Rather less enhancement was observed at the south polar cap in its summer because it

remains cold enough to retain H₂O as an ice and as a clathrate. A **clathrate compound** occurs when a substance, H₂O in this case, crystallizes with an open structure that can admit small molecules, here CO₂, into the cavities. These small molecules are held in the cavities until the crystalline substance is warmed or perhaps melted.

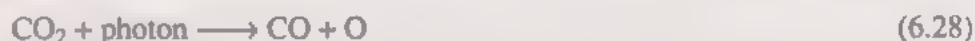
There is a continual seasonal exchange of CO₂ and H₂O between the polar caps via the atmosphere (Figure 6.20). Similarly, the soil-like surface exchanges CO₂ and H₂O with the atmosphere seasonally. The amount of these gases absorbed in the surface may be 10 to 100 times the amount in the atmosphere and, via the atmosphere, they exchange with the polar caps. Longer-term variations in the spin axis and orbit of Mars suggest that the atmospheric pressure may vary from about 10⁻³ bar to 2 × 10⁻² bar, possibly as a result of contributions from the absorbed gas.

Table 6.7 Composition of the troposphere of Mars

Gas	Volume ratio ^a	Major source	Major sink
CO ₂	9.53 × 10 ⁻¹	evaporation, outgassing	condensation
N ₂	2.7 × 10 ⁻²	outgassing	escape (as N)
⁴⁰ Ar	1.6 × 10 ⁻²	outgassing	
O ₂	1.3 × 10 ⁻³	photochemistry	photoreduction
CO	7 × 10 ⁻⁴	photochemistry	photo-oxidation
H ₂ O	3 × 10 ⁻⁴	evaporation, desorption	condensation, adsorption
³⁶ Ar	5 × 10 ⁻⁶	outgassing	
Ne	2.5 × 10 ⁻⁶	outgassing	
Kr	3 × 10 ⁻⁷	outgassing	
Xe	8 × 10 ⁻⁸	outgassing	photochemistry
O ₃	(0.1–20) × 10 ⁻⁸	photochemistry	photochemistry
NO	7 × 10 ⁻⁵ (120 km)	photochemistry	

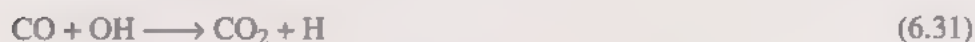
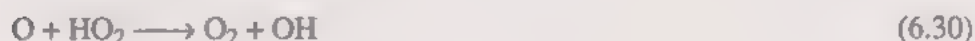
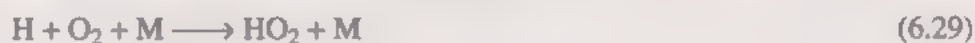
^a Values at the surface unless indicated otherwise.

Being very thin, the Martian atmosphere is completely exposed to solar UV radiation right to the surface. We therefore expect that, as on Venus, CO₂ will dissociate:



As on Venus, there is a problem of time-scale. In the absence of further reactions, the entire atmosphere of CO₂ would be destroyed in about 3 000 years.

The key to the chemistry that maintains CO₂ in the Martian atmosphere is the intervention of H₂O. Thus, recombination of CO and O takes place mainly in the lower atmosphere through the intervention of H, OH and HO₂ (hydroperoxyl) radicals. These radicals occur when H₂O is photodissociated. As in the Earth's troposphere, the OH radical acts to oxidize CO, that is it reacts to increase the proportion of oxygen in a compound (see Box 6.4). Starting with atomic hydrogen, a sequence of reactions leading to CO₂ can be deduced from laboratory studies:



The net effect of these three reactions is the conversion of CO and O into CO₂. It is just one of a number of schemes that can be devised using known chemical reactions of OH and HO₂ that are thought to be effective in regenerating CO₂ in the Martian atmosphere. Note that this sequence involves atomic oxygen, present from the dissociation of CO₂. Its reaction with molecular oxygen, O₂, accounts for the presence of small amounts of ozone, O₃, via Equation 6.19 (see the Chapman scheme).

Both O₃ and HO₂ are powerful oxidizing molecules. These and other oxidizing molecules have effects beyond interactions with atmospheric components, as follows.

Amongst the experiments on the Viking lander mission of 1976 were several tests designed to search for evidence of life on the planet. In one of these, nutrients such as amino acids, labelled with radioactive carbon (¹⁴C), were fed to a sample of the Martian soil, and the evolved gas was examined for the presence of labelled carbon dioxide, ¹⁴CO₂. If respiration had occurred, some ¹⁴CO₂ should have been released (see Equation 6.12). The first results of this experiment were exciting for biologists: ¹⁴CO₂ was detected in the evolved gas. But disappointment followed when further treatment of the same soil sample failed to yield any ¹⁴CO₂. Although this second result does not rule out bacteriological action by the soil, it does make it less likely.

A more plausible explanation of these results is now thought to be chemical action by the soil on the nutrient. The oxidizing substances most likely to be responsible for the CO₂ are peroxides, either H₂O₂ itself or some metal peroxide formed by interaction with surface rocks. It is certainly apparent that the surface of Mars is highly oxidized, the red colour (Plates 2.12–2.16, 2.31 and 2.32) indicating iron to be present in the Fe³⁺ state – this form of iron is in oxygen-rich iron compounds. The oxidizing atmosphere of the Earth also ensures that iron is in this form at the Earth's surface.

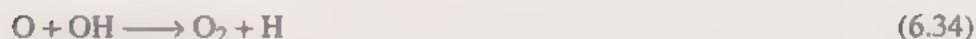
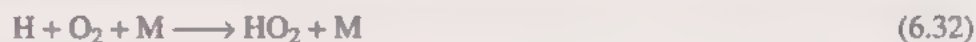
Further evidence for surface interactions has come from laboratory studies that were stimulated by the Viking biology experiments. Exposure of materials thought to be similar to Martian soil to UV radiation in the presence of a Martian-type atmosphere generates peroxides in or absorbed on the soil and, moreover, reproduces the results of the Viking lander experiments.

Summary of Section 6.5 and SAQs

- 1 The chemistry of the Earth's troposphere is dominated by biological processes that generate reduced forms of carbon. These molecules are eventually oxidized (chemically) by the action of hydroxyl radicals, and (biologically) by respiration and decay. Similarly, N₂ is 'fixed', converted into compounds of nitrogen, by life-forms, which are also responsible for the conversion back into N₂.
- 2 The chemistry of Venus divides into sub-cloud chemistry and processes above the cloud deck. The cloud material, H₂SO₄, is likely to be generated from SO₂, possibly by reactions involving NO and NO₂. Above the clouds the chemistry is dominated by the photodissociation of CO₂ and its reformation by reaction cycles, including those involving chlorine.
- 3 CO₂ is photodissociated throughout the Martian atmosphere. It is reformed in chemical cycles, probably involving O₂ and OH, which are very minor but active atmospheric components.

SAQ 6.7 (Objective 6.7) In the Chapman scheme, atomic oxygen, O, is effective in producing ozone by Reaction 6.19. Other species are also involved in

ozone chemistry, including the hydroperoxyl radical, HO_2 . What is the net effect of the following reactions?



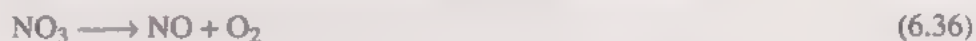
SAQ 6.8 (Objectives 6.1, 6.8 and 6.9) This SAQ is about the Earth's atmosphere.

(a) Nitrogen trioxide, NO_3 , is formed by the reaction of two minor atmospheric components (see Table 6.5):



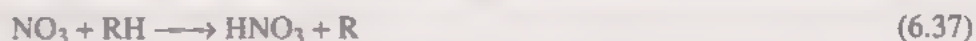
Why is a third body M *not* written in this reaction?

(b) NO_3 is dissociated either thermally or by UV radiation:



Would you predict Equation 6.36 to be exothermic or endothermic?

(c) NO_3 reacts rapidly with hydrocarbons, which we can represent by RH:



Subsequently, R reacts with O_2 and becomes oxidized. Overall, we can thus regard NO_3 as a radical that oxidizes hydrocarbons. In view of the ready photodissociation of NO_3 , when would you expect NO_3 to be unable to compete with OH in the oxidation process of hydrocarbons?

SAQ 6.9 (Objective 6.9) Why is the Martian surface at least as oxidized as the Earth's surface, although the Martian atmosphere (CO_2) is not as oxidizing as the Earth's atmosphere (N_2 and O_2)? Would you expect the surface of Venus to be as oxidized as that of Mars?

6.6 Origin and evolution

6.6.1 Origins of the atmospheres: some evidence

The atmospheres of Venus, Earth and Mars are quite different in composition and amount, yet the three planets are generally considered to have been formed by similar processes. It is reasonable to suppose that their atmospheres should have derived from similar sources. The variation in their present atmospheres presents a problem in trying to trace this origin. Moreover, the three planets, together with their atmospheres, were formed from the same material: the cloud of gas and dust that was the solar nebula (Section 2.2), and which consisted of about 98% by mass of hydrogen and helium. These two elements are only very minor components of the present atmospheres, which derive from the remaining 2% of the material of the solar nebula, from the somewhat heavier elements. Evidently, the terrestrial planets, when compared with solar or cosmic abundances, are highly depleted in hydrogen and helium.

When considering the terrestrial planets, it is important to recognize that a substance that exists in the atmosphere of one planet may be condensed or present in a combined form as a solid on another planet. As you saw in Section 6.4, the physical state of a substance is dependent on the prevailing conditions. On Mars, water exists mainly as ice, with very little as vapour in the atmosphere. The water on Earth is contained largely in the oceans, whereas Venus is so hot that water does not condense on the surface. In comparing the carbon dioxide contents, we must take account of such factors as the large sedimentary carbonate deposits of limestone on Earth, in which the CO_2 is now contained, as the carbonate ion, CO_3^{2-} . The relative abundances of volatiles of the terrestrial planets and of the Sun can then be estimated (Table 6.8). Such estimates are highly uncertain, however, because these planets might contain volatile substances at great depth.

Table 6.8 Observed or estimated relative volatile inventories of the Sun^a, Venus, Earth and Mars, expressed as proportions of the total mass

Volatile	Sun ^a	Venus	Earth	Mars
CO_2	1.5×10^{-2}	1.0×10^{-4}	1.0×10^{-4}	1.8×10^{-6}
H_2O	0.7	4×10^{-9}	4×10^{-4}	? ^c
N_2	9×10^{-4}	2×10^{-6}	2.4×10^{-4}	$\sim 10^{-9}$
F	4×10^{-7}	2×10^{-13} ^b	1×10^{-6}	?
S	3.6×10^{-4}	5×10^{-3}	6×10^{-6}	4×10^{-2}
Cl	8.1×10^{-6}	3×10^{-3}	6×10^{-6}	8×10^{-3}
^{20}Ne	2.2×10^{-3}	2.9×10^{-10}	6×10^{-11}	4.4×10^{-14}
^{36}Ar	9.0×10^{-5}	2.5×10^{-9}	3.5×10^{-11}	2.2×10^{-10}
^{40}Ar	$< 9 \times 10^{-5}$ ^c	2.5×10^{-9}	1.0×10^{-8}	6.2×10^{-10}

^a The abundances in the Sun column are based on the formula of the volatile and are not meant to suggest that these volatiles exist in the Sun. Thus the CO_2 abundance is based on the C content of the Sun and H_2O is based on the H content. This is not meant to imply that the O:H ratio is 1:2 as in the formula H_2O . The Sun contains *much* less oxygen than hydrogen.

^b This is an estimate of the atmospheric abundance.

^c Values for these are uncertain

FIGURE 6.9 Table 6.8 is intended to include those volatile elements near the Earth's surface, but not those in the deep interior. Estimate the total mass of H_2O that would be present in the Earth's mantle if it contained 0.1% of H_2O by mass. Take the thickness of the mantle to be 3000 km and its average density to be $4.0 \times 10^3 \text{ kg m}^{-3}$.

The data in Table 6.8 are therefore the known amounts, which exclude volatiles contained in the deep interiors of the planets. Notice that the forms of the volatiles listed in Table 6.8 are those that are prevalent at the surfaces. Thus hydrogen occurs largely as H_2O , carbon as CO_2 , and nitrogen as N_2 . In the Sun, of course, these elements exist as ions in a plasma, but for comparison their abundances are converted into the equivalent amounts of H_2O , CO_2 and N_2 .

From the data in Table 6.8, does it appear that the terrestrial planets derived their volatiles from the same source as did the Sun?

■ The large differences in relative amounts of volatiles between these planets and the Sun suggests a different source, but of course all of this volatile material originally existed in some form in the solar nebula.

It has been estimated that substantial atmospheres of H_2 and He were captured by the terrestrial planets from the solar nebula, about 3×10^5 times the mass of the present atmosphere in the case of the Earth (Subsection 2.5.3). The lack of H_2 and He in the present atmospheres suggests that they did not originate from the direct capture of gas from the solar nebula, but are derived from some secondary origin.

Nor does it appear that these atmospheres are some small remnant of atmospheres that were captured directly as gas and have since lost almost all of their hydrogen and helium. Evidence against the notion of small remnant atmospheres is found by comparing the C, N and H abundances with those of the noble gases that are not produced during radioactive decay, the non-radiogenic noble gases: ^{20}Ne and ^{36}Ar .

How do the depletions (relative to the Sun) of these noble gases compare with those of the other volatiles, C, N and H, in Table 6.8?

■ In general for Earth and Mars, the non-radiogenic gases are more depleted than C, N and H. For example, in the case of the Earth, C is depleted relative to the Sun by a factor of 150, whereas ^{36}Ar is depleted by a factor of 2.6×10^6 . For Venus, ^{20}Ne and ^{36}Ar are more depleted than are C and N. (As we shall see, there are special reasons for the apparent depletion of H_2O on Venus.)

If C, N, H, ^{20}Ne and ^{36}Ar were only the remnants of a more massive atmosphere from which most of the light gases H_2 and He had been lost, their depletions would be more similar.

The implication here is that C, N and H are less depleted because they can be chemically combined (whereas the noble gases are rather inert to chemical reaction), and are derived from material in the solar nebula in which they were rendered temporarily less volatile, by chemical combination, than the noble gases.

This is consistent with the idea that the initial substantial atmosphere that was captured from the solar nebula has been lost. In this case, the terrestrial planets now possess volatiles that were gained in some less volatile form at the formation of these planets and have subsequently been released by outgassing (see Subsection 6.6.3). Further evidence for this notion is the enrichment in radiogenic noble gases, such as ^{40}Ar (Table 6.8), which are produced in rocks by the decay of radioactive elements, such as ^{40}K , and then outgassed. We next examine how the original atmospheres may have been lost, before returning to the question of how the present volatile material was acquired.

6.6.2 Loss of volatiles

First we consider why there is so little H_2 and He in the terrestrial planets and why their atmospheres are thought not to be the result of the direct capture of gas.

According to most theories of the formation of the Solar System, the planets accreted within the cloud of gas and dust that was the solar nebula. By gravitational attraction each planet would then collect an atmosphere from the surrounding gases, composed mainly of H_2 and He.

The retention of such an atmosphere would depend on the balance between the gravitational attraction by the planet and the tendency of fast molecules to escape from the top of the atmosphere. The escape speed, v_{esc} (see *Preparatory science*, Section 1.5), is determined by the properties of the planet

$$v_{\text{esc}} = \left(\frac{2GM}{R} \right)^{1/2} \quad (6.38)$$

where M is the mass of the planet, and R is the radius of the planet. If the molecules at the top of the atmosphere are in thermal equilibrium, and we shall

assume that this is so in spite of infrequent collisions between them, their root mean square speed (*Preparatory science*, Subsection 2.4.1) is

$$v_{\text{rms}} = \left(\frac{3kT}{m} \right)^{1/2} \quad (6.39)$$

where m is the molecular mass. According to Equation 6.39, it is the lightest molecules that achieve escape speed most easily. For retention of a gas over the present lifetime of the Solar System, v_{rms} should be less than about one sixth of v_{esc} . Figure 1.10 shows that, over this enormous interval, the Earth and Mars would lose hydrogen and helium, while Venus should retain both gases. However, there is geological evidence that the Earth had already lost these gases when the Solar System was a tiny fraction of its present age. Therefore, the estimate of the rate of escape of H_2 and He is too low, had existing conditions prevailed, to account for the supposed loss of the primitive atmospheres in a relatively short period. So, escape in which the speed of the gas molecules depends on the temperature (so-called **thermal escape**) does not satisfactorily explain the loss.

How might different conditions in the early evolution of the Solar System have contributed to this loss?

■ You will recall from Subsection 2.5.3 that before joining the main sequence, stars evolve through the T Tauri phase, characterized by powerful stellar winds.

At this stage, the Sun probably had a greater luminosity than it has today, and in particular a greater UV flux. According to the accepted model, the effect of both the stellar wind and the UV radiation is to impart greater speeds to the particles in the upper atmosphere which is exposed to the radiation. Collisions between particles of the solar wind (mainly protons) result in energy transfer to atmospheric components. UV radiation, in the wavelength range 30–85 nm, is absorbed by hydrogen molecules, with the result that the molecule dissociates into two atoms with high kinetic energy. These atoms can escape directly or transfer their kinetic energy, in collisions, to other particles, e.g. He atoms, providing them with sufficient speed to escape. Additionally, the general effect of the UV radiation and the solar wind is to raise the temperature of the upper atmosphere, so facilitating thermal escape. For a star of solar mass, the T Tauri phase extends over about 2×10^7 years, which is considered to be sufficiently long to sustain the loss of the entire primitive atmospheres of the terrestrial planets. At the same time, the gas escaping from the planets is driven from the inner Solar System, so that these planets do not recapture H_2 and He .

6.6.3 Gain of volatiles

One terrestrial planet, Mercury, stands out from the others in the composition of its atmosphere and in the way it has acquired it. Its atmosphere is very thin and consists of helium and hydrogen, which it gains by capture from the (now gentle) solar wind. This is evidenced by the fact that the heavier element, helium, is much in excess. This capture is only temporary, however, because the gravitational attraction by this small planet is weak and its temperature is high. There is thus a balance between gravitational capture and loss by thermal escape. Recent radar evidence indicates that Mercury may have patches of water ice permanently in shadowed regions close to the poles, probably supplied by outgassing and meteorites.

The other three terrestrial planets, Venus, Earth and Mars, seem to have gained their volatiles not directly as gas but in some less volatile form. The most widely accepted theory is that the volatile material was present in the dust grains from which the planets accreted, via planetesimals. These grains supposedly contained entrapped gases and other volatile elements that were chemically

combined within the grains. The presence of these volatiles in such material is shown by the volatile content of the carbonaceous chondrites (Figure 6.30), which, as you saw in Subsections 2.2.3 and 2.3.2, are believed to be relatively unprocessed meteorites that also accreted from dust grains in the solar nebula. Support for the grain accretion hypothesis comes from the similarity in the relative amounts of volatiles (see Figure 6.30), especially noble gases, between these meteorites and the three planets. The dissimilarity with the pattern of abundances for the Sun underlines the argument made earlier against the direct capture of gas from the solar nebula.

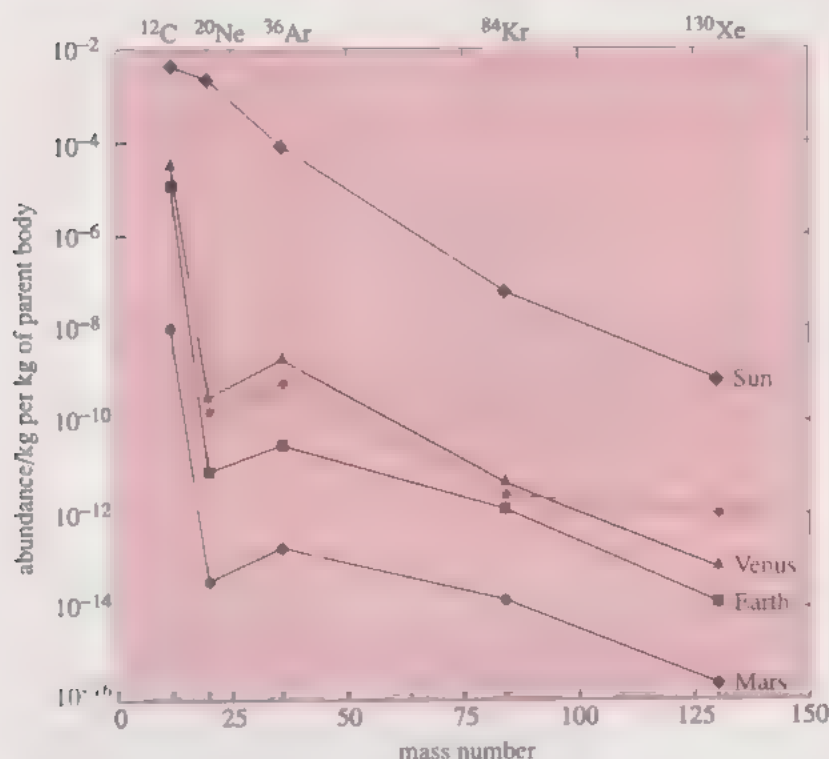


Figure 6.30 A comparison of the abundances of carbon and the non-radiogenic noble gases in the Sun, the terrestrial planets and carbonaceous chondrites (CC). The pattern of relative abundances for the terrestrial planets suggests a common origin similar to carbonaceous chondrites and unlike that of the Sun. These values differ somewhat from those in Table 6.8 – different scientists come up with different estimates!

The stripping of the primitive atmosphere during the Sun's T Tauri phase left the terrestrial planets without atmospheres. Thus deprived, the terrestrial planets are thought to have gained atmospheres from the **outgassing** of material that was contained in the solid material that accreted (grains to planetary embryos) to form these planets. Evidence for outgassing is to be found in the relative abundances of the **radiogenic noble gases**. To take one example, the isotope ^{40}K of the involatile element potassium undergoes radioactive decay to the isotope ^{40}Ar . Geological activity, such as volcanism, brings this gaseous element to the surface, where it becomes part of the atmosphere. The subsequent enrichment of the atmospheres of the terrestrial planets in ^{40}Ar , compared with the non-radiogenic isotope ^{36}Ar , is shown by the data in Table 6.8. Other sources, such as capture of gas from the solar wind and gain of volatiles from meteorites and comets, have probably made only minor contributions, except for Mercury.

The outgassing process is thought to have made its major contribution to the atmosphere in the period 4.5 to 3.5×10^9 years ago, the outer regions of the terrestrial planets being hot enough to outgas from the earliest times. During this outgassing, the chemical reaction of metal oxides and silicates with the material of carbonaceous chondrites, which contains C, N and H, released the gases CO_2 , N_2 and H_2O . The unreactive noble gases, which were trapped with the grains,

were also released largely during this early period, probably mostly during the lifetime of a magma ocean (Subsection 3.3.2). Volcanism added more later.

Various factors have determined how much outgassing occurred during this period. The temperature gradient in the solar nebula would have resulted in a gradient of volatiles outward from the Sun. The volatile components incorporated into the terrestrial planets would also have been affected by the timing of accretion, as the gaseous component of the solar nebula was swept away by the proto-Sun. The extent of outgassing may also have been different for Venus, Earth and Mars, and there is some evidence for this. In view of these different factors, it is difficult to reach firm conclusions about the total amounts of volatiles expected to have been produced by outgassing of these planets.

6.6.4 Evolution of the atmospheres

We have concluded that three of the four terrestrial planets possess atmospheres that have a common origin. Yet the compositions of these atmospheres are now quite different from each other, suggesting that the three planets have evolved in different ways. A starting point for tracing the evolution of atmospheres that derived from outgassing could be the composition of gases released by volcanoes on Earth. A typical composition is given in Table 6.9. On this basis, we might expect the atmospheres initially to contain much H₂O and CO₂, together with N₂ and other minor components.

However, current terrestrial outgassing may not represent the composition of the initial outgassing of the Earth or other planets. In particular, the volatiles trapped in the material that accreted to form planetary embryos would be determined largely by the local temperature in the early solar nebula, which in turn would depend on distance from the proto-Sun. Thus, volcanism on Earth might not be a good model for other planets. Additionally, plate tectonic activity on Earth effectively recycles the volatile materials. So there may be significant differences between the emission of current volcanoes and that of early outgassing. However, in the absence of better information, volcanic emission gives the best indication of that outgassing.

Currently, the atmosphere and surface of Venus appear to possess much less H₂O than Earth. Can it be that Venus was formed with very little H₂O, and is it sensible to expect similar volatile inventories in these two planets?

- What evidence exists in the form of other volatiles to indicate that the two planets once had rather similar inventories of volatiles?
- Venus and Earth contain similar amounts of other volatiles, N₂ and CO₂ (Table 6.8), although on Earth the CO₂ is mainly contained in carbonate deposits (see Section 6.3).

Thus the initial inventories of Venus and Earth are likely to have been similar, although differences cannot be ruled out on this evidence alone. Here, we shall suppose that there were no major differences in the initial volatile inventories of the three planets, although estimates of the temperature of the solar nebula suggest that there must have been minor differences. On this basis, we shall examine how the three planets may have achieved their present atmospheres from an origin that resembled current terrestrial volcanic emissions.

6.6.5 The atmosphere of Venus

Comparison of the present atmospheric composition of Venus (Table 6.6) with the gases produced by outgassing on Earth (Table 6.9) reveals one striking difference already mentioned. Venus has almost no H₂O in its atmosphere. Moreover, the surface temperature of about 730 K rules out the existence of reservoirs of liquid water on the planet. If, as we suppose, Venus formed with a substantial amount of H₂O, it must have retained its water in its interior or lost

Table 6.9 Gases from active Hawaiian volcanoes

Constituent	Percentage by number of molecules
H ₂ O	73.5
CO ₂	11.8
N ₂	4.7
SO ₂	6.6
SO ₃	2.3
CO	0.5
H ₂	0.4
Cl ₂	0.05
Ar	0.2

almost all of this volatile substance while retaining its other volatiles, CO₂, N₂ and other minor components. It can be argued that the mantle of Venus contains hydrous minerals, although there is no direct evidence for this notion. However, we shall consider how loss of H₂O may have occurred and examine the evidence that supports this idea.

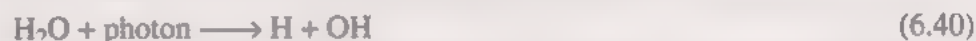
Loss of H₂O from the planet requires that H₂O should have been present in the atmosphere rather than condensed in oceans, as on Earth. Venus is hotter than Earth because it is nearer the Sun, and for another reason.

□ What other factor ensures a higher temperature?

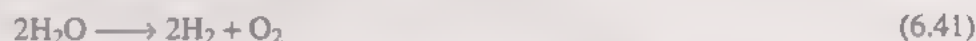
■ The high surface temperature of Venus at present is due largely to the greenhouse effect (Subsection 6.4.2), which might have been operative for most of the lifetime of Venus.

Even if the surface temperature of Venus had initially favoured the condensation of H₂O, its closeness to the Sun means that the partial pressure of H₂O would have been sufficiently high for H₂O to make its own contribution to the greenhouse effect. In fact, H₂O absorbs radiation in different parts of the spectrum from CO₂ (Figure 6.7), so its contribution is important in these circumstances. Any rise in surface temperature resulting from this greenhouse effect would ensure the evaporation of more liquid H₂O. The increased partial pressure of H₂O in the atmosphere would then further contribute to the greenhouse effect. Such a scenario is an example of positive feedback and is termed, for obvious reasons, the **runaway greenhouse effect**. It would eventually lead to the complete evaporation of any liquid H₂O on Venus, on a time-scale that depends on the rate of outgassing of the planet. Rapid outgassing, perhaps over less than 10⁸ years, would have resulted in the formation of oceans, owing to the high initial partial pressure of H₂O coupled with a relatively low initial temperature. Other estimates set the period of outgassing as long as 10⁹ years, when it is less likely that water would have liquefied.

In either case, the partial pressure of H₂O in the atmosphere would have been high, of the order of 100 bar in rapid outgassing and somewhat less than 10 bar in slow outgassing. The fate of H₂O in the atmosphere, as of many other substances exposed to solar radiation, is photodissociation:



Atomic hydrogen and the hydroxyl radical are reactive species, which ultimately convert H₂O into molecular hydrogen and oxygen. The net effect of this conversion is summarized as:



The high UV flux of the early Sun would thus have efficiently converted H₂O into these products.

□ By which process are gases lost by atmospheres?

■ You should recall (Subsection 6.6.2) that thermal escape is possible for sufficiently light molecules.

The effect of solar UV heating and the solar wind on the upper atmosphere of Venus at this time would have been sufficient to enable H₂ to escape, although escape of O₂ would have been much less likely on account of its higher molecular mass. Other processes were available for the removal of O₂ from the atmosphere. Reaction with the surface is the most plausible, provided that a substantial area of surface had been resurfaced by volcanism or processed by plate recycling, in order to remove the oxidized products and expose fresh material for reaction. In fact, the Venera 13 lander observed the presence of oxidized iron (Fe³⁺) minerals on the surface, which is evidence for this process of

O₂ removal. The eventual outcome of this whole process is a planet whose atmosphere and lithosphere are both dry.

Some evidence for this proposed loss of H₂O is to be found in the results of the analysis of the atmosphere by the Pioneer spacecraft. The spacecraft examined the relative amounts of hydrogen and deuterium, an isotope of hydrogen, which has twice the atomic mass of hydrogen. The results showed that there is relatively much more deuterium on Venus than on Earth. This enrichment indicates that a lot of H₂O has been lost from Venus, possibly as much as there is at present on Earth. Loss of hydrogen, ¹H, occurs preferentially to loss of deuterium, ²H, because of the greater mass of deuterium.

Loss of H₂O leaves Venus with an atmosphere that is substantially CO₂, with some N₂ and noble gases, as observed (Table 6.6).

6.6.6 *The atmosphere of Mars*

For Mars, the property that sets it apart from Venus and Earth is its low volatile inventory – see Table 6.8 and (from different estimates) Figure 6.30.

Which factors could be responsible for the low volatile inventory of Mars?

■ If, as is supposed, the atmosphere and surface volatiles derived from outgassing, it could be that Mars originally possessed a smaller amount of volatile material, or that it has not outgassed to the same extent as the other planets.

The temperature gradient in the solar nebula would mitigate against a low original inventory, so that less complete outgassing seems more likely. If this is the case, the Martian atmosphere has never been as extensive as that of Venus or Earth. There is, however, evidence of at least one former period of more substantial atmosphere than exists at present. Currently, the atmospheric pressure is so low that liquid water does not exist, the transition from gas to solid occurring directly. Images of the surface of Mars show features that resemble canyons and valleys, whereas others look distinctly like channels on Earth (Figure 6.31, Plate 2.15), indications that water once flowed as a liquid on Mars. The cratering time-scale (Subsection 4.4.4) of the regions that contain these features has been used to estimate when these valleys were formed. Although different interpretations have been made of this evidence, the valleys and gullies were not formed during a single, early event. Channels were formed during the period of late heavy bombardment, but there is also evidence of water flow about 3 billion years ago.

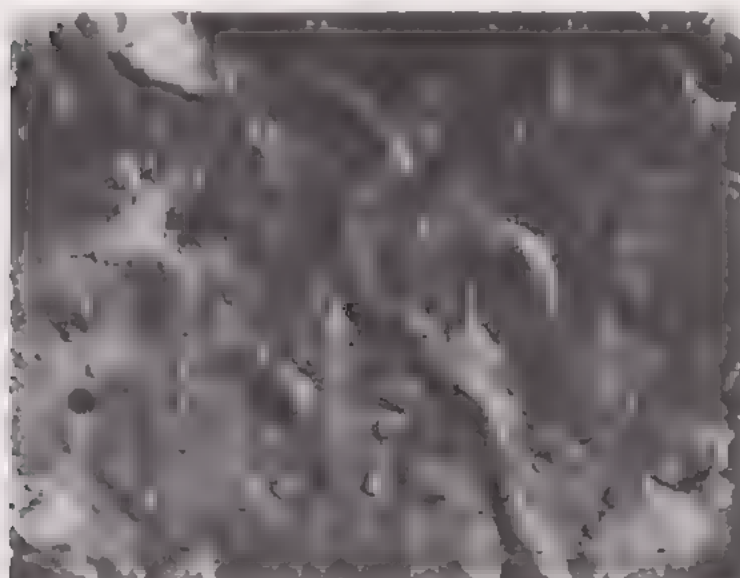


Figure 6.31 An image of the Martian surface, showing features that resemble channels cut by flowing water on Earth. The area shown is about 320 km across.

The most plausible explanation for these features is that the surface temperature of Mars has been higher than it is at present. A higher temperature would result in the evaporation of some of the condensed volatiles, for example H_2O and especially CO_2 . Mars would then have possessed a more substantial atmosphere. Subsequent loss of some of this atmosphere would have led to a cooling of the planet, because the surface temperature is strongly determined by the bulk of the atmosphere. Any loss of water or other atmospheric components would have occurred not only by thermal escape but also by various other processes. Consequently, the remaining atmosphere should carry the signature of this loss in the enrichment of heavier isotopes. Of particular interest in this context is the isotopic analysis of volatiles from the surface of Mars as measured in certain meteorites for which there is compelling evidence of Martian origin, as you will see in Section 8.5. It is in that chapter that we examine the isotopic evidence of past climatic conditions on Mars.

In summary, the evidence of atmospheric evolution on Mars points to a planet in which outgassing is less complete than on Venus and Earth, and in which the atmosphere has been lost partly to space (Chapter 8) and then to the surface as the temperature consequently fell.

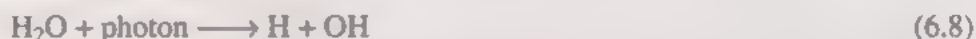
6.6.7 *The atmosphere of Earth*

At present, the Earth's atmosphere is dominated by the influence of life. Many of its components are produced in significant amounts only by biological processes and are destroyed by biological and chemical processes. It is supposed that the early atmosphere bore some resemblance to the composition of current volcanic gases. Two major questions that arise concern the formation of oceans from H_2O that entered the atmosphere as gas, and the removal of CO_2 together with the production of O_2 by plant life. On Earth, we have access to the geological record, and so the evidence for evolution of the atmosphere is more soundly based than for other planets.

The geological record indicates that Earth has possessed oceans of liquid water since early times. The Isua deposits of Western Greenland, which are dated at 3.8×10^9 years old, are some of the oldest known rocks on Earth, and contain sedimentary deposits. Their formation shows that CO_2 was being removed from the atmosphere at that time, by dissolution in water and precipitation as insoluble metal carbonates (mainly calcium carbonate). Studies of these rocks indicate that the Earth's temperature was close to its present-day value, which is surprising in view of the major change that the atmosphere and the Sun have subsequently undergone.

It is also known, from chemistry, that the early atmosphere was mildly reducing. By reducing, in this context, we mean that the volatile elements, such as carbon and nitrogen, were combined with hydrogen rather than oxygen. So compounds such as CH_4 and NH_3 would have been present, although not as the most abundant carbon-containing or nitrogen-containing molecules.

In the UV radiation of the early Sun, the photodissociation of H_2O generated OH radicals:



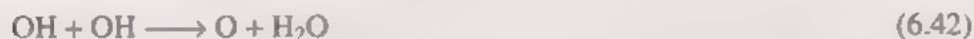
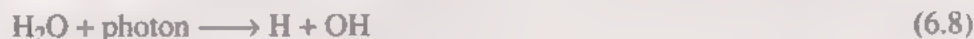
What is the fate of reducing substances, such as CH_4 and NH_3 , in such circumstances?

■ These compounds are oxidized by OH (see Subsection 6.5.2).

Small amounts of H_2 , released by volcanoes, would have been present, but in the presence of OH (because of the fast reaction of OH with H_2) its abundance would have been limited.

From the geological record, it also appears that the prebiotic atmosphere, far from being strongly reducing, contained O_2 , although only in small amounts. In rocks older than 2.5×10^9 years, the minerals pyrite (FeS_2) and uraninite (U_3O_8) have been found. These minerals are readily oxidized, the products dissolving in water, so an upper limit of about 2×10^{-4} bar can be set on the partial pressure of O_2 . Simultaneously, oxidized iron (as Fe_3O_4) is present in rocks, requiring at least some O_2 in the atmosphere ($> 10^{-11}$ bar).

This oxygen of the early atmosphere could have readily resulted from the photodissociation of H_2O or CO_2 :



Estimates of the production of O_2 by photodissociation during a period of intense solar UV flux suggest a much lower maximum for the partial pressure of about 5×10^{-9} bar.

The prebiotic atmosphere therefore consisted mainly of CO_2 , H_2O and N_2 , with small amounts of both oxidizing (O_2) and reducing (H_2 , CH_4 and NH_3) gases and other minor substances, such as CO. It was the emergence of life that transformed this atmosphere. Again it is the geological record that provides a time-scale for this evolution. Although the earliest large organisms with easily fossilized hard parts were not formed until about the start of the Cambrian geological period, about 550 million years ago, far older rocks, dating to more than 3.5×10^9 years ago, contain evidence of bacteria and other microfossils, including some in the form of algal colonies known as stromatolites. These bear a striking resemblance to present-day algal colonies that exist in highly saline environments (see Figure 6.32). At that time the lethal solar UV radiation would have penetrated to the Earth's surface, and all life-forms would have had to be aquatic for the protection offered by the absorbing layer of water. Photosynthesis by these early life-forms would have raised the O_2 level of the atmosphere, as carbon became incorporated into biological matter (see Equation 6.12).

The term 'prebiotic' refers to the atmosphere before its composition had been altered by the influence of life.

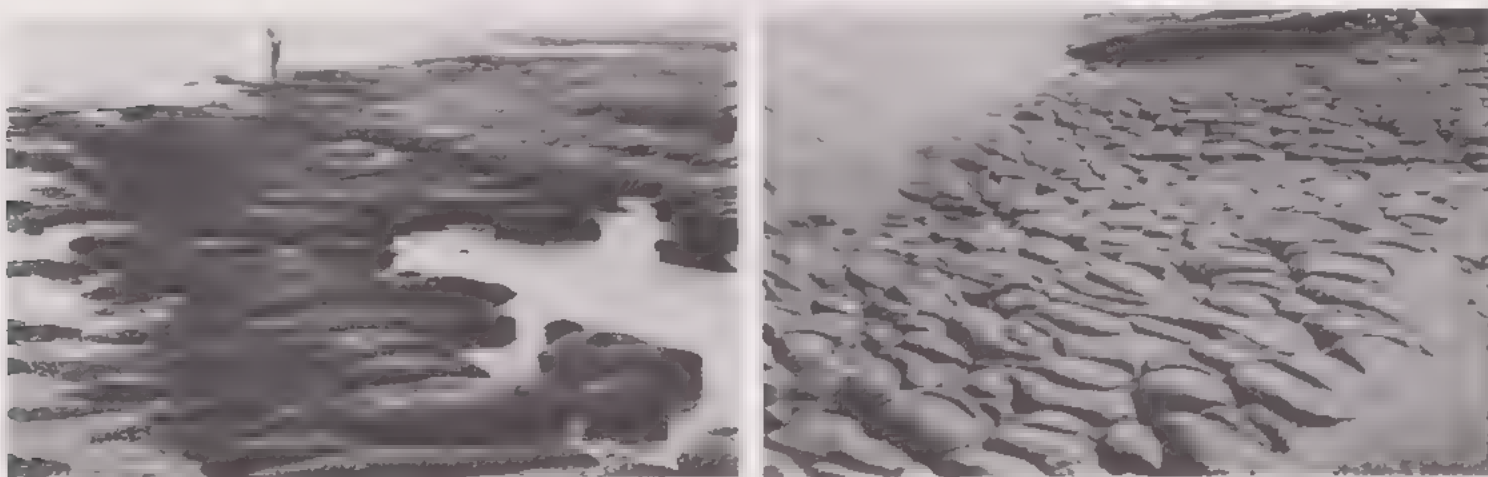


Figure 6.32 Ancient stromatolites (right) resemble present-day algal colonies.

The fossil record contains signs of the continuous build-up of O_2 in the atmosphere. About 2×10^9 years ago there emerged a different type of organism based on cells containing a nucleus, the eukaryotic cells. (Cells lacking a nucleus are called prokaryotic.) Eukaryotic cells require a partial pressure of O_2 of about 10^{-3} bar in order to function. At about 600 million years ago, soft-bodied metazoan fossils (organisms like jellyfish) appear in the geological record. These

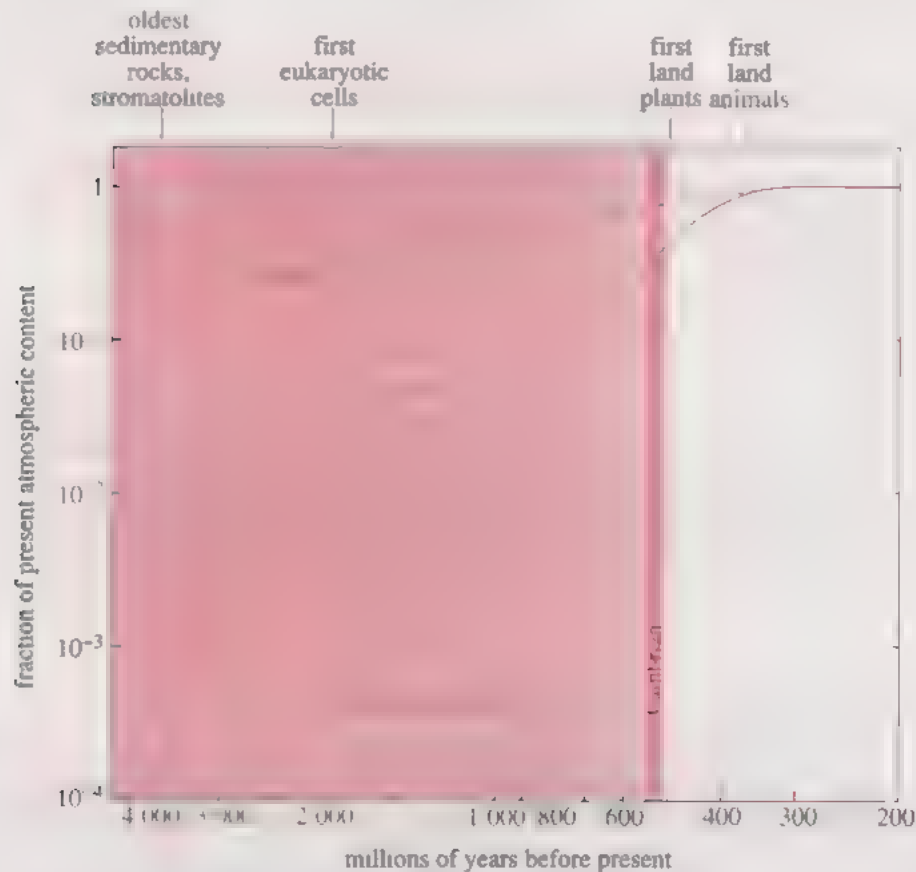
creatures absorb oxygen through their surface and require a partial pressure of O_2 of about 2×10^{-2} bar. Finally, by the end of the Silurian geological period, about 410 million years ago, land animals appeared, by which time the O_2 concentration would have reached a value similar to that of today.

Concurrent with the increase of O_2 , we can conclude that ozone (O_3) also built up in the stratosphere.

□ What effect would this have on the evolution of life?

■ As O_3 increased, the UV flux at the surface would diminish, enabling life-forms to exist in shallower water and eventually to emerge on to land.

Ozone concentrations can be estimated only by modelling the effects of solar radiation interacting with the atmosphere. The results of one such estimate of O_2 and O_3 variations are shown in Figure 6.33, which indicates that for several hundred million years the atmosphere has been stable.



Cambrian is the name of the geological period near the beginning of which metazoan animals with readily fossilizable hard-parts first became abundant.

Figure 6.33 A model of planetary evolution shows the increase of oxygen, O_2 , and ozone, O_3 , as life colonized the Earth. The more marked increases in O_2 content correspond roughly to the early evolution of eukaryotic marine life-forms, such as algae, and the colonization of the land by plants when the build-up of ozone offered protection from UV radiation.

A major current concern is the disturbance of the stability of our atmosphere by human activities. From studies such as that illustrated by Figure 6.33, it is reasonable to suppose that without such activity our atmosphere would continue to be stable for a long period. But human activities have been changing the atmosphere since the Industrial Revolution. In particular, the level of CO_2 is rising steadily, as measurements in several locations over many years have shown. For example, Figure 6.34 gives the variation of atmospheric CO_2 since 1958 at Hawaii. Notice that the sawtooth pattern arises from seasonal variation linked to the growth of plant life in summer by photosynthesis, but that this is superimposed on a steady increase. Estimates of earlier levels of CO_2 , made by

analysis of bubbles trapped in ice-cores from Antarctica, show that for thousands of years prior to the Industrial Revolution the CO_2 level had been steady at 260 parts per million (260×10^{-6} volume ratio).

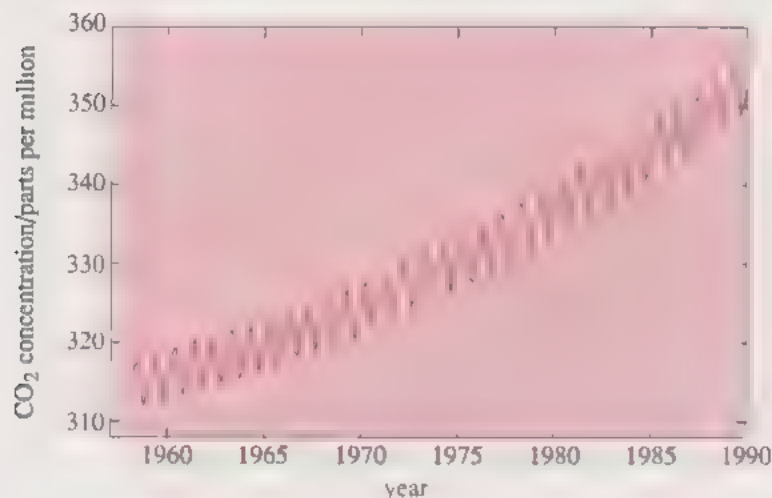


Figure 6.34 The variation of CO_2 content of the atmosphere as recorded at Mauna Loa, Hawaii, since 1958.

What effect will increased CO_2 have on the temperature of the Earth's surface?

■ CO_2 contributes to the greenhouse effect and so will tend to raise the temperature.

Other gases arising from industry and agriculture are also greenhouse gases, that is they contribute to the greenhouse effect. These include the chlorofluorocarbons and their replacements (which are designed to have no effect on the ozone layer), as well as methane, which is generated by agriculture, particularly paddy fields.

A warmer Earth would affect the H_2O reservoirs, resulting in some melting of polar ice-caps, with release of entrapped gases, as well as increased evaporation from the oceans, thereby raising the H_2O level in the atmosphere. Now H_2O is itself a greenhouse gas and will contribute further to the rise in temperature. The possibility exists that such a positive feedback could result in a runaway greenhouse effect. Such is the concern about this possibility, it has even caught the attention of politicians.

Summary of Section 6.6 and SAQs

- 1 The lack of H_2 and He in the atmospheres of the terrestrial planets, together with the pattern of abundances of non-radiogenic noble gases, indicates that the present atmospheric gases were not captured directly from the solar nebula.
- 2 Any early atmosphere of H_2 and He would have escaped from the planets by the combined effect of the solar wind and UV radiation during the T Tauri phase of the Sun.
- 3 The secondary atmospheres of the planets are very probably the products of outgassing, the most likely original source being grains, of similar composition to carbonaceous chondrites, that constituted the material from which the terrestrial planets formed.
- 4 The dry atmosphere of Venus suggests that its H_2O , present initially in its atmosphere and hydrosphere, has been lost. This suggestion is supported by the enrichment of deuterium. Such loss may have occurred through

photodissociation of H_2O , with escape of H to space and incorporation of O into the surface.

- 5 On Mars, geological features indicating the flow of water suggest that Mars once had a warmer, denser atmosphere. It appears that H_2O from the atmosphere has become trapped as polar ice-caps and permafrost, and some has escaped from the planet.
- 6 The Earth's early atmosphere is supposed to have consisted mainly of CO_2 , H_2O and N_2 . The earliest rocks contain evidence of the existence of oceans that removed much CO_2 from the atmosphere to form carbonate rocks. Fossils also date from the earliest rocks. The fossil record of evolving life, together with the evidence of reduced and oxidized minerals, enables us to chart the increase of O_2 in the Earth's atmosphere due to the colonization of the planet by life. Current changes in the atmosphere have human origins.

SAQ 6.10 (Objectives 6.10 and 6.11) The planet Mercury has a very thin atmosphere, which consists mainly of He with a little H_2 . If its source is the solar wind (of solar composition), why should He dominate rather than H_2 ?

SAQ 6.11 (Objectives 6.1, 6.11 and 6.13) The relative amounts of the radiogenic isotope ^{40}Ar on Venus, Earth and Mars are 3×10^{-9} , 10×10^{-9} and 0.6×10^{-9} (kg of ^{40}Ar per kg of planet). If the non volatile components of the nebula from which the planets formed were roughly the same for these three planets, what do you conclude about the relative degrees of outgassing? (Recall that ^{40}Ar is the product of decay of ^{40}K .)

SAQ 6.12 (Objectives 6.1 and 6.12) Conditions on Venus before the onset of the runaway greenhouse effect may have been quite similar to those on Earth. What evidence would you require to prove the former existence of oceans on Venus, and how might it be collected?

SAQ 6.13 (Objective 6.13) It has been suggested that the transport to Mars of sufficient greenhouse gases would eventually make the planet able to support life. How would you expect the greenhouse gases to transform the atmosphere?

SAQ 6.14 (Objectives 6.11 and 6.12) Why was the presence of liquid water on Earth essential for the formation of O_2 in the atmosphere?

Objectives for Chapter 6

After studying Chapter 6 (and any associated audio, video or TV material), you should be able to:

- 6.1 Give brief definitions of the terms, concepts and principles listed at the end of the Objectives.
- 6.2 Evaluate various analytical techniques for specific purposes in atmospheric studies.
- 6.3 Relate the albedo of a body to its effective temperature and its likely surface coverage.
- 6.4 Account for the effect of infrared-absorbing gases on the temperature of an atmosphere.
- 6.5 Predict the formation of solid or liquid material condensing from planetary atmospheres using a saturation vapour pressure diagram.

- 6.6 Describe the various forms of atmospheric motion (relates also to video sequence 7).
- 6.7 Determine the net chemical effect of a given scheme of reactions.
- 6.8 Describe the general requirements and outcome of the chemical reactions involving dissociation and combination.
- 6.9 Describe the chemical and physical effects of atmospheric chemistry, and of interactions between the surface and the atmosphere.
- 6.10 Predict or explain the effects of thermal escape on a planet's atmosphere.
- 6.11 Use evidence of atmospheric composition to draw conclusions about the origins of the atmospheres.
- 6.12 Cite evidence to support the supposed evolution of a planet's atmosphere.
- 6.13 Deduce evolutionary atmospheric changes from given evidence.

List of scientific terms, concepts and principles used in Chapter 6

Term	Page	Term	Page	Term	Page
albedo	126	gas chromatography	123	reduced substance	145
atmospheric structure	127	greenhouse effect	129	reducing substance	145
bond formation	142	Hadley cell	134	reduction	145
carbon cycle	146	heteronuclear molecule	123	relative molecular mass	123
Chapman scheme	147	homolytic dissociation	141	rotational spectroscopy	122
chemical energy	140	homonuclear molecule	123	runaway greenhouse effect	160
chemical equilibrium	141	mass spectrometry	123	saturation vapour pressure diagram	132
chemical reactions	139	mesosphere	131	seasons	136
clathrate compound	152	nitrogen fixation	146	selection rules	123
collision rate	141	noble gases	122	sink	143
column mass	126	outgassing	158	steady state	139
condensation flow	137	oxidation	145	stratosphere	131
Coriolis effect	135	oxidized substance	145	super-rotation	136
diatomic molecule	122	oxidizing substance	145	thermal escape	157
dissociation	141	ozone	131	thermal tide	137
diurnal variation	124	partial pressure	131	thermosphere	130
effective temperature	127	photochemistry	142	third body	142
electric dipole	122	photodissociation	142	three-body collision	142
electron pair bond	141	photosynthesis	143	troposphere	130
endothermic reaction	140	radical	141	vibrational spectroscopy	122
energy barrier	140	radiogenic noble gases	158	volatile inventory	127
exothermic reaction	141	rate of reaction	141		

Chapter 7

The Giant Planets

Prepared for the Course Team by Elaine Moore

Contents

7.1	Introduction	169
7.2	The interiors of the giant planets	170
7.2.1	Jupiter and Saturn	173
7.2.2	Uranus and Neptune	176
	Summary of Section 7.2 and SAQs	177
7.3	Composition of the atmospheres	178
7.3.1	The atmospheres of the giant planets	178
7.3.2	The atmosphere of Titan	181
	Summary of Section 7.3 and SAQs	182
7.4	How the atmospheres vary with depth	182
7.4.1	The giant planets	182
7.4.2	Titan	186
	Summary of Section 7.4 and SAQ	187
7.5	Atmospheres in motion	187
7.5.1	Jupiter	189
7.5.2	The Great Red Spot	190
7.5.3	Winds on Saturn	191
7.5.4	The atmospheric circulation of Neptune	191
7.5.5	Neptune's Dark Spot	193
7.5.6	Winds in the atmosphere of Uranus	194
	Summary of Section 7.5 and SAQs	195
7.6	Atmospheric chemistry	196
7.6.1	Nitrogen on Titan	196
7.6.2	Photochemistry of hydrocarbons	198
	Summary Section 7.6 and SAQs	200
7.7	Overview of the giant planets	200
	Objectives for Chapter 7	201
	Answers to Exercise on Neptune's Dark Spot	202

7.1 Introduction

Most of this Book so far has dealt with objects with definite surfaces – the terrestrial planets and planetary satellites. We turn now to objects that are more reminiscent of stars, in that there is no discernible surface and that the greater part of the object (and possibly all) is fluid (that is gaseous or liquid). These are the giant planets, Jupiter, Saturn, Uranus and Neptune. Jupiter, indeed, has sometimes been compared to a small star with its own planetary system of satellites. There are some resemblances to the pattern of the Solar System in Jupiter and its satellites, as you saw in Chapter 2 (Section 2.6), but there are also substantial differences and we would not wish to push this analogy too far.

We start by considering the overall structure of these planets. In Chapter 2 (Subsection 2.5.3), you saw that they are thought to consist mainly of hydrogen and helium captured by rocky/icy kernels. Here we shall look at the observational evidence for their structure. Much of the detailed evidence has come from spacecraft, and it is hoped that even more information will be gleaned by future missions (see the table in *Images of the Cosmos*). However, Earth-based instruments are by no means obsolete in this field. Ground-based observations were necessary as a starting point for data collection by the spacecraft, and these and observations by the Hubble Space Telescope (in orbit round the Earth) continue to add to our knowledge today (Plates 2.34 and 2.42). For example, Earth-based observations can study changes in a planet's appearance over a long time (in the case of Jupiter hundreds of years), whereas spacecraft observe for only a limited time. Even with all the data collected, however, there is still room for uncertainty; the results may not give a definite answer but be capable of explanation by one or more models. We shall see, for example, that the data on the interiors of Uranus and Neptune in particular are still open to various interpretations.

Our knowledge of the composition and structure of the outer layers or atmospheres of these planets is more certain, because we can detect and positively identify molecules in them. As none of the four planets has an accessible surface (if they have one at all), we do not know where the base of the atmosphere is. The radii of the planets are defined as the distance from the centre of the planet to the 1 bar pressure level (1 bar being approximately the pressure of the Earth's atmosphere at sea-level). This level corresponds roughly to the base of the directly observable layers of the planets.

For an atmosphere, we can determine not only its composition and structure but also how it moves. Are the processes determining this motion the same as those on the terrestrial planets, or does the far greater depth of fluid on the giant planets mean that different processes occur? Can the way gas is transported around the atmosphere tell us anything of what goes on in the interior of the planet? We shall be looking at the observed wind motions and seeing if we can draw any conclusions from them. Finally in this chapter you will study some of the chemical processes occurring in the atmospheres. The composition of the atmospheres is very different from that of the atmospheres of the terrestrial planets, and so we might expect different chemical reactions to occur. The largest satellite of Saturn and the only satellite with a substantial atmosphere, Titan, is included here because its atmospheric chemistry is more similar to that of the giant planets than to that of Earth.

There are few new scientific concepts in this chapter, but concepts you have met earlier in the Course will be applied to the (often very different) conditions found in the giant planets and Titan. Because of the difficulties in collecting data from these bodies, you will often find that there is no one clear-cut explanation and that several plausible models will fit the data that we do have. Consequently, there are some sections in which you will not be asked to learn an accepted explanation but to consider whether a particular model is ruled out by the

information available. Do not worry, therefore, if you leave this chapter feeling you are not quite sure, for example, what the internal structure of Uranus is. Nobody is sure!

7.2 The interiors of the giant planets

Figure 7.1 shows sketches of the interiors of Jupiter, Saturn, Uranus and Neptune. These represent one interpretation of the data currently available.

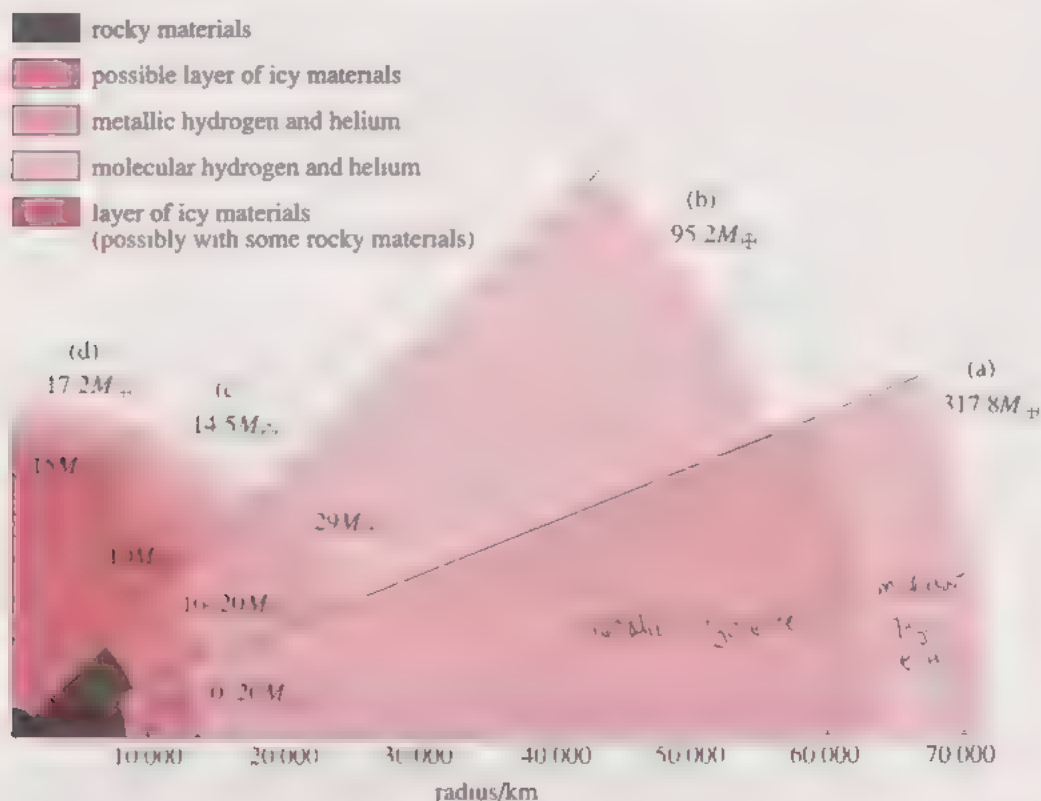


Figure 7.1 Cross-sections through (a) Jupiter, (b) Saturn, (c) Uranus and (d) Neptune, showing the various layers that are thought to exist in the interiors of these planets. The icy materials are *not* cold!

Note that the regions are shown by volume and that this does not represent the mass ratios of the different regions. The cumulative masses of the layers are labelled on Figure 7.1.

Some of the layers marked may strike you as odd: what, for example, is metallic hydrogen? On Earth, hydrogen is a colourless gas that is very difficult to liquefy and is certainly not metallic.

Let's consider the evidence on which such models are based. The first and incontrovertible piece of evidence is the average density of each planet as a whole. The diameters of the planets can be measured and so their volumes can be found.

Saturn has an equatorial radius of 60 000 km. If Saturn is spherical, what is its volume?

■ The volume of a sphere is $\frac{4}{3}\pi R^3$, where R is the radius. Thus the volume of Saturn is $V = \frac{4}{3} \times \pi \times (60\,000)^3 \text{ km}^3 = 9.0 \times 10^{14} \text{ km}^3 = 9.0 \times 10^{23} \text{ m}^3$.

The planets are not quite spherical, being flattened at the poles. However, the volume can be calculated if the polar and equatorial radii are measured. The other datum needed to calculate the density is the planetary mass. The masses of all four planets have been obtained by studying the paths of their satellites and of spacecraft flying past. These smaller bodies are affected by the gravitational

attraction between them and the planet, which of course depends on the planetary mass.

ITQ 7.1 The period of a satellite of Saturn, Enceladus, is 1.370 days and its average distance from Saturn is 238 000 km. Assuming that the mass of Saturn is much greater than that of Enceladus and that the satellite's orbit is roughly circular, estimate the mass of Saturn using the formula $M = 4\pi^2 \frac{a^3}{G\tau^2}$ (Subsection 1.2.2).

ITQ 7.2 Now estimate the average density of Saturn.

The densities obtained from the volumes and masses are given in Table 1.1a, where accurate volumes, allowing for the polar flattening, have been used.

The densities of the giant planets are very low. The mass of material in these planets is so large that the interior must be at a very high pressure: in the case of Jupiter, the pressure increases towards the centre from Earth's atmospheric pressure at the edge of the atmosphere to 80 million times this at the centre of the planet. With such high pressures, the only way to account for the low overall density is that the composition of these planets is mainly the light elements, hydrogen and helium, in some form. Saturn is less massive than Jupiter.

☐ Saturn is also less dense than Jupiter. Does it follow that the two planets differ in composition?

☒ Not necessarily. The greater mass of Jupiter means that the material will be subject to greater self-compression and therefore be more dense. It is probable in fact that Jupiter and Saturn have similar chemical compositions.

Uranus and Neptune are considerably smaller than Saturn so that their higher densities tell us that they have a higher ratio of heavier elements to hydrogen and helium.

The passage of spacecraft and satellites tells us more than just the mass of the planet. Because the planets are not spherical, it turns out that from the gravitational field we can also learn something of the distribution of density in the planet. For example, we know that there are denser and lighter materials and that the denser materials lie near the centre as in Figure 7.1, rather than being distributed throughout the planet as in Figure 7.2.

Incidentally, this method is also the best way of measuring the internal density changes of Venus and Mars in the absence of seismic data (Subsection 3.2.3). The measurements we obtain do not give us a definitive model for the nature of the interior, but any model proposed must be able to reproduce the observed results. It is possible, for example, to use this method to place limits on the ratio of hydrogen to helium in the envelopes surrounding the cores of Jupiter and Saturn.

Another source of information on the interiors of the planets is measurements on the magnetic fields. Like the Earth, the four giant planets have a magnetic field that behaves as though it originated in a magnetic dipole.

Can you recall from Chapter 3 what properties the Earth's core had to have to explain this magnetic dipole field?

☒ It had to be liquid, electrically conducting, and in motion (Subsection 3.2.2).

Jupiter's magnetic field has been known and studied for many years, but Voyager 2 produced the first comprehensive measurements on the magnetic fields of

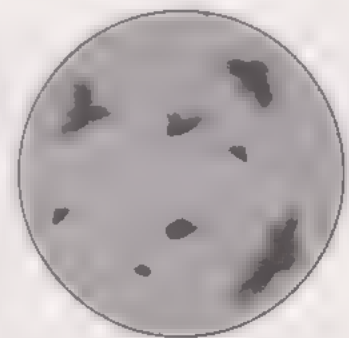


Figure 7.2 Hypothetical (and unrealistic) random distribution of dense material in a planet.

Uranus and Neptune. The existence of the magnetic fields means that somewhere in the interior of each planet lies an electrically conducting liquid.

What is the electrically conducting liquid inside the Earth?

■ It is the outer core, which is liquid iron with about 4% nickel and about 10% of a lighter element (Subsection 3.2.3).

In the four giant planets, a different fluid is thought to be responsible. Extensive mapping of the magnetic fields has allowed the region in which the magnetic field originates to be determined. Figure 7.3 shows the main features of the magnetic fields. The 'bar magnets' representing a dipole will reproduce the observed fields.

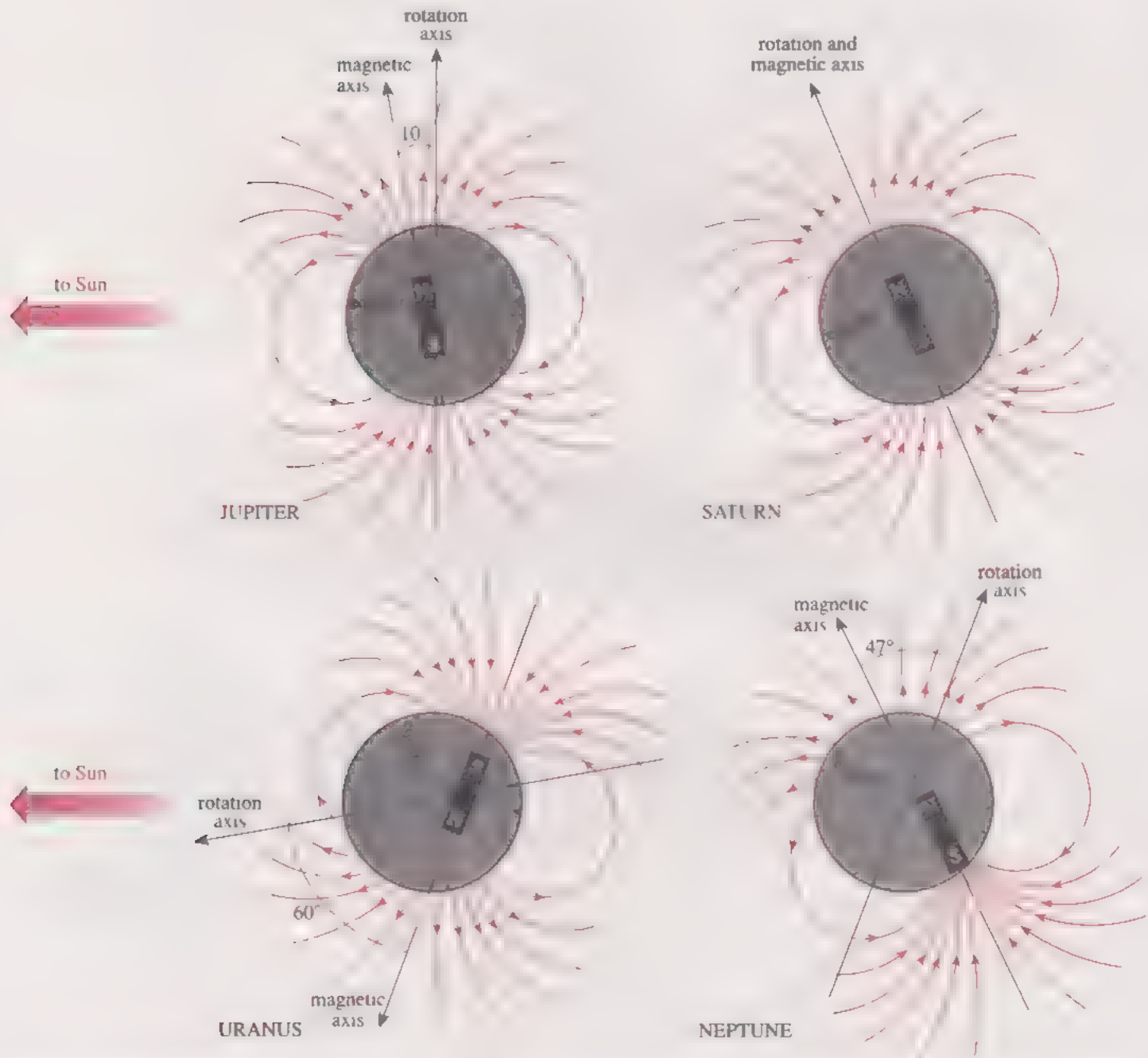


Figure 7.3 Magnetic dipoles of Jupiter, Saturn, Uranus and Neptune. The 'bar magnets' indicate the position of the centre of the dipole and its polarity. The length is arbitrary. The lines represent observed magnetic field lines. The direction to the Sun is at one point in the orbit: the directions of the rotation and magnetic axes are fixed with respect to the stars, not the Sun (see Figure 6.19).

Detailed measurements on Jupiter and Saturn show that the magnetic source is outside the core in the layer labelled metallic hydrogen in Figure 7.1. The region where the magnetism originates is a spherical shell in this layer, and the magnetic dipole produced by such a shell is centred in the planetary core.

For Uranus and Neptune, the magnetic source is a long way out from the centre in the layer of icy materials. The best dipole approximation to these sources places the dipoles off centre and in the southern hemispheres.

For our purposes, the main conclusion from such measurements is that there must be an electrically conducting liquid layer in the metallic hydrogen regions of Jupiter and Saturn, or in the icy materials region of Uranus and Neptune.

More evidence on the nature of the interiors comes from the heat given out by the planets. The effective temperatures of the planets are very low: 124 K (Jupiter), 95 K (Saturn), 59 K (Uranus) and 59 K (Neptune). They are, however, except for Uranus, higher than would be expected from their distance from the Sun, if the Sun were their only source of heat. Jupiter, Saturn and Neptune must therefore have an internal source of heat, and it is possible to estimate how hot the interior is from the amount of heat radiated by the planet. In the case of Jupiter, the core temperature is thought to be over 20 000 K.

All these indications of internal layering within the giant planets support the view that they underwent differentiation subsequent to their formation. It would almost certainly be wrong to equate their present cores with the kernels around which most of the gas was captured during planetary formation, as described in Subsection 2.5.3, although a kernel would obviously be a good starting point in core formation.

Finally, in the absence of evidence to the contrary, it is assumed that heavy elements are present in roughly solar ratios and that the planets are more likely to contain substances made from more abundant elements than less abundant ones. On these grounds, for example, just as in the icy satellites (Section 2.6), ammonia (NH_3) is a more likely major constituent of icy materials than hydrogen fluoride (HF).

7.2.1 Jupiter and Saturn

From density measurements it is inferred that Jupiter and Saturn each contain about 10–20 Earth masses of icy and rocky material. Measurements of the gravitational fields of these planets are consistent with most of this being in the central core. Models of the planets that fit these measurements, and agree with our current knowledge of how the density of materials varies with temperature and pressure, show a core of dense material. In the absence of further information, we can reasonably assume that this core is composed of common rocky and icy planetary materials, such as silicates, iron compounds, water, ammonia and methane. In some models there is a central rocky core surrounded by icy material, as in Figure 7.1; in others there is just a rocky core, the icy materials being dispersed in the hydrogen/helium layers.

Surrounding the core is a very thick layer taking up the greater part of the planets, labelled metallic hydrogen and helium. The helium is believed to be present as a fluid, similar to that used on Earth in very low temperature studies, but at much greater temperatures and pressures. But what is meant by **metallic hydrogen**? To see how hydrogen can occur in a form described as metallic, read Box 7.1.

Box 7.1 Metallic bonding

A simple picture of a metal such as aluminium, familiar to us on Earth, is of an orderly set of positively charged ions surrounded by electrons, which are shared. Imagine some aluminium ions arranged as in Figure 7.4.

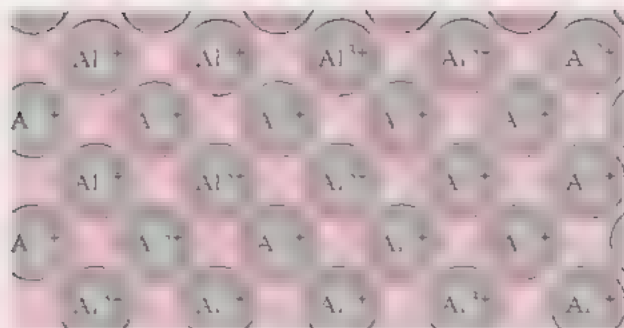


Figure 7.4 Aluminium ions arranged as though in a crystal of aluminium. The electrons 'lost' to form the ions wander freely through the solid.

Now, the most common ion formed by aluminium is Al^{3+} , in which the aluminium atom has lost three electrons. Suppose the lost electrons are allowed to wander freely about the crystal. You may recall that two atoms can form a chemical bond by sharing electrons. In our picture of aluminium, the freely wandering electrons are shared by all the ions in the crystal, and thus serve to bond together all the ions. This sharing of electrons by a whole crystal constitutes **metallic bonding**. The electrons are responsible for many of the characteristic metallic properties of elements such as aluminium. In particular, they can move through the solid and, because they are electrons, they carry negative electrical charge with them. A moving electrical charge is an electric current, and so metallic bonding leads to electrical conductivity.

Aluminium is a solid at everyday temperatures and pressures, but there is one common liquid metal found at the Earth's surface. What is it?

■ Mercury (Hg) You may have a mercury thermometer or barometer at home.

In liquid mercury, the situation is similar to that shown in Figure 7.4 except that the ions are not so ordered (Figure 7.5).

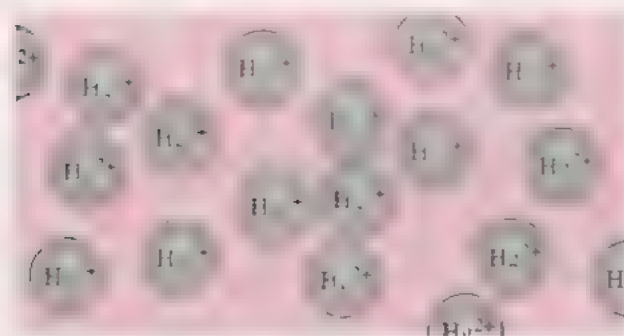


Figure 7.5 Mercury ions arranged as though in liquid mercury. The 'lost' electrons are free to travel through the liquid as in aluminium, but the arrangement of the ions is less regular.

What is thought to happen at the very high pressures in the interior of Jupiter and Saturn is that the hydrogen nuclei are squashed together. Then, instead of just two hydrogen nuclei sharing electrons to form H_2 (which is the normal form of the element hydrogen on Earth), large numbers of hydrogen nuclei share their electrons, and **metallic bonding** results as in mercury.

Metallic hydrogen is liquid at the temperatures and pressures inside these planets. If metallic hydrogen is indeed the form in which hydrogen is present in

the layers indicated in Figure 7.1, then this provides a ready explanation for the magnetic field. The electrically conducting fluid is liquid metallic hydrogen. The heat of the interior will cause the liquid to be in motion through, for example, convection currents.

As we move out from the core, however, temperatures and pressures are reached where hydrogen is more stable as H_2 molecules than as a metal. This provides the outer layer of the interior. As we rise through the outer layer, the hydrogen changes imperceptibly from liquid to gas.

For Jupiter, the model that we have described accounts for the measurements of gravity, magnetic field and internal energy. For Saturn, we have to introduce a further refinement to account for the internal energy. Jupiter is large enough for the internal heat to be residual accretional heat and heat due to past differentiation of rocky and icy materials (Subsection 3.5.2) but Saturn should have cooled down long ago.

☐ What is the main heat source for the interior of the Earth?

☒ Radiogenic heating (Subsection 3.5.2).

The bulk of Saturn, however, is hydrogen and helium, and density measurements preclude sufficient abundances of radioactive isotopes for this to be a major form of heating.

ITQ 7.3 Jupiter generates a heat flux of about 7 watts per square metre of surface at the 1 bar radius. The decay of ^{40}K and other radiogenic isotopes produces $4.8 \times 10^{-9} \text{ mW kg}^{-1}$ of heat in the Earth today. If Jupiter has a core of which $8M_{\oplus}$ is rocky materials, what would be the rate of radiogenic heat generation in the core of Jupiter, assuming that the core is of similar composition to the Earth? Compare this with the observed heat flux.

The extra heat in Saturn may come from the separation of hydrogen and helium in the metallic hydrogen layer. In Jupiter, this layer is hotter and well stirred by convection. At the higher temperatures the two liquids mix to form a solution. You are probably familiar with two liquids such as alcohol and water that can be mixed to form one liquid (such as wine). When other liquids such as olive oil and water are added together they form two layers. The situation at the lower temperatures on Saturn is rather like the oil and water but there are not yet the separate layers. If you stir the two liquids vigorously you can get them mixed (as for example in salad dressing) but if the stirring becomes less vigorous, droplets of one liquid form in the other and gradually sink to the bottom to form a separate layer. Convection stirs the layers in Saturn so that the hydrogen and helium are still in the process of un-mixing. The liquid helium in Saturn's metallic hydrogen layer is thought to be forming droplets which, because they are denser than the hydrogen, fall towards the centre of the planet and will eventually form a separate layer.

This is comparable with the tequila sunrise example that you met in the answer to ITQ 3.4.

☐ How does this separation cause the interior of Saturn to heat up?

☒ This is an example of heat of differentiation. The helium droplets as they fall release gravitational energy, which is converted into heat.

7.2.2 *Uranus and Neptune*

These two planets are believed to be very similar to each other but different from Jupiter and Saturn. They are denser than Jupiter and Saturn so must have a lower proportion of hydrogen and helium to heavier nuclei.

If Uranus and Neptune had a similar elemental composition to Jupiter and Saturn, how would you expect their densities to compare?

■ The densities of Uranus and Neptune would be much lower because the smaller sizes of these planets would result in less self-compression.

The theory of planetary formation in Subsection 2.5.3 suggests that Uranus and Neptune began life with rock/ice kernels, but there is no direct evidence for differentiated rocky cores today. Measurements on the gravitational field show that the density increases towards the centre, but this could be due to the increasing self-compression of icy materials. Most models indicate that complete separation of icy and rocky materials has not occurred. The boundary between the icy layer and the hydrogen/helium layer is probably not sharp, and icy materials are probably found in the outer layers.

The magnetic fields of Uranus and Neptune originate in a magnetic dipole source that is some way from the centre of the planet and whose axis is considerably tipped with respect to the spin axis (Figure 7.3). When it was found that the magnetic axis of Uranus was at 60° to its spin axis, it was suggested that Voyager may have caught the planet as its magnetic field was changing its polarity. A change of magnetic field polarity is not unknown in planetary science: the magnetic field of the Earth is known to have changed polarity many times.

When Neptune's magnetic field was also found to be tilted, this time at 47° to the spin axis, then it was felt there must be some other explanation as the coincidence of two planets simultaneously undergoing a change of polarity was too unlikely. Some calculations indicate that, if the electrically conducting shell is thin ($0.1R_U$ or $0.2R_N$) then the magnetic axis has to be tilted.

The position of the magnetic dipole source (roughly one-third of the way out from the centre for Uranus and just over half-way out for Neptune) puts it well outside any rocky core that may be there, in a layer believed to be mostly liquid icy materials.

A mixture dominated by water, ammonia and methane does not look a very promising candidate for an electrically conducting layer. However, experiments on water under high pressure and extrapolation to the temperatures and pressures found in Uranus and Neptune suggest that there will be much greater ionization of water in these layers than in the more familiar surroundings of the Earth's surface. In addition, ammonia dissolved in water forms NH_4^+ ions. The most popular view, therefore, is that the currents giving rise to the magnetic fields of Uranus and Neptune are carried by the ions NH_4^+ , H_3O^+ , and OH^- .

The final line of evidence is the internal heat of the planets. Uranus and Neptune have almost identical effective temperatures, despite Neptune's greater distance from the Sun. These temperature measurements lead to the conclusion that Neptune emits more heat than it absorbs (as do Jupiter and Saturn) but that Uranus does not.

One possible explanation is that Neptune has a hot interior and Uranus a cold one, but this is thought unlikely, in view of the similarities in composition and size. The currently accepted explanation is that both interiors are hot, but that in Uranus this heat is somehow prevented from reaching the outer edges and escaping. Mechanisms have been proposed to account for this, but there is no entirely satisfactory explanation for this difference between Uranus and Neptune. The internal heat is thought to be due to heat of differentiation. Several schemes of differentiation can be suggested, but it is difficult to decide which one is most plausible because of the uncertainties in the composition of the interior.

Summary of Section 7.2 and SAQs

- 1 Data on the interiors of the giant planets can be obtained from measurements of gravitational field, magnetic field and emitted heat.
- 2 Jupiter and Saturn are believed to contain 10–20 Earth masses of rocky and icy materials, most of which is in their cores. These cores are surrounded by layers that are mostly hydrogen and helium, which account for most of the planets' mass.
- 3 Uranus and Neptune probably also have rocky cores, but surrounding the core is a mantle of mainly icy materials and around this is a layer of mainly hydrogen and helium. Overall, these two planets are less dominated by hydrogen and helium than Jupiter and Saturn.
- 4 The magnetic fields of Jupiter and Saturn are believed to originate in a shell of liquid metallic hydrogen.
- 5 The magnetic fields of Uranus and Neptune are thought to originate in a shell of liquid icy material containing the ions H_3O^+ , OH^- and NH_4^+ .
- 6 Jupiter emits more energy than it receives from the Sun (heat excess). This is thought to be due to the continuing escape of original accretional heat and heat of differentiation.
- 7 Saturn's heat excess is thought to have an additional contribution from helium droplets separating out from metallic hydrogen and sinking.
- 8 The deep interiors of both Jupiter and Saturn are very hot (over 20 000 K in the case of Jupiter).
- 9 Neptune has a net emission of energy, but Uranus has not. It is thought both planets have hot interiors, but the causes of the heat excess for Neptune and the lack of it for Uranus are still debatable.
- 10 Jupiter and Saturn do not have a definite liquid or solid surface.
- 11 Uranus and Neptune may or may not have a surface. The data are not sufficient to say with any certainty.

SAQ 7.1 (Objectives 7.2 and 7.3) The average density of Jupiter is very similar to that of the icy satellite of Saturn, Rhea. Rhea is thought to be composed of half rock and half icy materials. How is the same average density compatible with very different chemical compositions of the two bodies?

SAQ 7.2 (Objectives 7.2 and 7.3) Outline the reasons why the element mercury is not a suitable candidate for the material responsible for producing the magnetic dipole in the four giant planets.

SAQ 7.3 (Objective 7.3) Before the Voyager mission reached Uranus and Neptune, it was suggested that Uranus had a hot, molten, rocky core, which gave rise to a magnetic field, but that the core in Neptune was solid and so Neptune would have no magnetic field. How far does this model fit the current observational data? (Consider magnetic field measurements and other measurements.)

SAQ 7.4 (Objectives 7.2 and 7.3) Laboratory work on a mixture of icy materials at very high pressures (around 2×10^6 bar) has led to a suggestion that Uranus and Neptune do not have rocky cores at all but are composed entirely of icy materials plus hydrogen and helium. Does such a suggestion conflict with either (a) the observational evidence or (b) the theory of planetary formation outlined in Chapter 2?

SAQ 7.5 (Objective 7.3) Measurements on the gravitational field of Saturn indicate that the hydrogen/helium ratio in the molecular hydrogen layer is roughly solar and that helium is not depleted. Does this fit in with the attribution of the heat excess of Saturn to differentiation of helium from metallic hydrogen?

7.3 Composition of the atmospheres

7.3.1 The atmospheres of the giant planets

The images we have of the giant planets, as shown for example in Plates 2.34–2.42, are really images of the outermost cloud layers. Sometimes there are holes in the cloud layer and we can see through to the next layer, and if we are very lucky we may glimpse even deeper into the atmosphere. By using different wavelengths of radiation we can study the atmosphere underneath the clouds, but even so identification of atoms and molecules in the atmosphere is limited to a layer that is very thin compared with the size of the planets.

The atmospheres of the four planets were studied at radio, infrared, visible and ultraviolet wavelengths by instruments on the Voyager probes. Some molecules have also been identified by Earth-based telescopes. Because of the low temperatures of the atmospheres, most elements are present as molecules rather than atoms, and most of the molecules are identified from their vibrational spectra.

Vibrational spectroscopy was discussed in Section 6.2 and in Chapter 5 of Book 1. What wavelength radiation is usually associated with vibrational spectra?

■ Infrared.

It has been possible to detect molecules with relatively low abundances, such as ethyne (acetylene) and phosphine, through their absorption of infrared radiation.

What do you suppose is the source of the infrared radiation absorbed?

■ The planets themselves. These act approximately as black-body emitters from layers at pressures of about 1 bar, and because the atmospheric temperatures are so low, the peak of the black-body radiation curve lies in the infrared.

The more abundant ammonia and methane were first detected through their absorption of visible sunlight reflected by the planetary atmosphere and received by telescopes on Earth. Methane and ammonia are colourless gases, and their electronic spectra lie in the ultraviolet. How then can we observe them by their absorption of visible light?

In vibrational spectra in the laboratory, the lines are usually produced by each molecule going from the lowest vibrational energy level to the next highest (Book 1, Subsection 5.2.5). It is possible for the molecules to go from the lowest level to higher levels, as in Figure 7.6, but this happens less often. If there is a very large number of molecules in the path of the radiation, as there are in planetary atmospheres, then there are enough molecules jumping several vibrational energy levels for the corresponding spectral lines to be observed. The absorption of red light by methane in the atmosphere of Uranus due to jumps of this sort is responsible for the blue-green tinge of the planet (as seen in Plate 2.34).

So far, we have not discussed the detection of the most abundant chemical species in the atmospheres, hydrogen (as H_2) and helium.

Can you spot a difficulty in detecting these elements through their vibrational spectra?

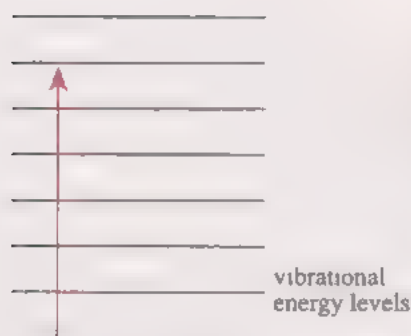


Figure 7.6 A vibrational transition to a high vibrational energy level, as observed for Uranus and Neptune.

■ Helium occurs as atoms, and so has no bonds to vibrate and no vibrational spectrum. Molecular hydrogen has vibrational energy levels but, as you saw in Section 6.2, homonuclear diatomic molecules, such as H_2 , have only an extremely weak vibrational spectrum.

Hydrogen can be detected in the ultraviolet through its electronic spectrum, but the detection of helium is a major problem. Its spectrum is in a part of the ultraviolet not normally covered by the instruments used and in which it is difficult to work. Helium abundance is therefore measured indirectly. Two methods are used, both of which depend on assuming that the major component of the atmosphere not accounted for must be helium atoms.

The first method is to look at the shape of the lines in the vibrational spectrum of, say, methane (CH_4). One factor affecting the width of the lines is the pressure of the atmosphere. If we have direct measurements of the abundance of the major constituents of the atmosphere other than helium, then we can calculate the effect these would have on the line shape. Assuming the rest of the atmosphere is helium, we can then calculate what pressure of helium is necessary to give the observed line shape.

Why do we assume the rest of the atmosphere is helium rather than some other atom or molecule with no suitable spectrum?

■ In the absence of evidence to the contrary, we assume roughly solar abundances, and helium is the second most abundant element in the Sun.

The second method involves sending radio waves through the atmospheres from a spacecraft and detecting them on Earth. Just as a ray of light travelling from air to water bends as it enters the water, so the radio waves are bent by the planetary atmosphere (Figure 7.7).

The amount by which the beam is deflected depends (among other things) on the average molecular mass of the gas mixture. Again, we assume that the undetected component is helium, and calculate how much helium is needed to give the average molecular mass measured. The identification of the missing constituent as helium is strengthened by the close agreement of the abundance of helium calculated by the two different methods.

The chemical compositions of the four atmospheres are given in Table 7.1

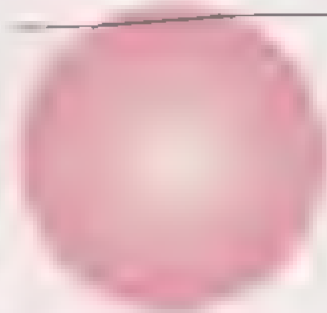


Figure 7.7 The path of radio waves through a planetary atmosphere. The depth of atmosphere penetrated by the radio waves is exaggerated to show the deflection.

Table 7.1 Chemical composition of the atmospheres of the giant planets

Atom or molecule	Abundance/fraction of total number of molecules ^a			
	Jupiter	Saturn	Uranus	Neptune
H_2	0.90	0.96	0.85	0.85
He	0.10	0.04	0.15	0.15
CH_4 (methane)	1.75×10^{-3}	1.76×10^{-3}	$<10^{-7}$	3×10^{-5}
NH_3 (ammonia)	1.8×10^{-4}	2×10^{-4}	—	—
H_2O	$1\text{--}30 \times 10^{-6}$	—	—	—
C_2H_6 (ethane)	5×10^{-6}	4.8×10^{-6}	—	3×10^{-5}
C_2H_2 (ethyne)	2×10^{-8}	1.1×10^{-7}	$3 \times 10^{-8}\text{--}2 \times 10^{-7}$	3×10^{-7}
PH_3 (phosphine)	6×10^{-7}	2×10^{-7}	—	—
CO (carbon monoxide)	$1\text{--}10 \times 10^{-9}$	—	—	—
GeH_4 (germane)	7×10^{-10}	—	—	—

^a A dash indicates that the atom or molecule has not been detected. The number of significant figures given reflects the uncertainty, so that a number quoted as 2 implies that the quantity lies between 1 and 3, whereas 2.0 implies that it lies between 1.9 and 2.1.

The presence of rocky/icy materials deeper down means that, overall, the giant planets have greater-than-solar abundances of elements heavier than hydrogen and helium.

There are several things to note that are relevant to theories of planetary evolution, to our models of the interiors of the planets, and to the chemical reactions occurring in the atmospheres.

The most obvious point is that all four atmospheres are predominantly hydrogen and helium, with a roughly solar ratio of hydrogen to helium. The predicted proportion of helium, assuming solar abundances, is 0.15. The atmosphere of Saturn has a low abundance of helium. The differentiation of helium from hydrogen, which was put forward as an explanation of the internal heat of Saturn (SAQ 7.5), could be responsible for this – if helium is separating out in the metallic hydrogen layer, then it will be replaced by helium from the outer layers.

The abundances listed in Table 7.1 were measured for the gaseous atmospheres above the cloud layers.

Why do the abundances given in Table 7.1 not represent the true abundances of molecules that are also found in the cloud layers?

■ The tabulated abundances are too small for compounds that are present as solid or liquid in the clouds – only the gas phase abundance is measured.

The outermost cloud layers on Uranus and Neptune are thought to be composed of methane, but the temperatures on Jupiter and Saturn are too high for methane to condense at the atmospheric pressures of these planets. Do the abundances in Table 7.1 reflect this?

■ Yes. The abundance of methane is much higher on Jupiter and Saturn than on Uranus and Neptune.

As you will see later, the abundances of ethane (C_2H_6) and ethyne (C_2H_2) can be explained only by looking at the chemical reactions in the atmosphere.

An important point to note is the types of molecule found. In the atmospheres of the terrestrial planets, carbon, for example, is found mainly as carbon dioxide, CO_2 . The carbon is joined to oxygen and is said to be in an oxidized form.

In the molecules listed in Table 7.1, what is carbon joined to?

■ Mostly hydrogen (CH_4 , C_2H_6 , C_2H_2). There is also a small amount of CO, in which the carbon is joined to oxygen but to only one atom of oxygen not two as in CO_2 .

Carbon in these molecules is in a reduced form.

Are the other elements detected present in reduced forms?

■ Yes. Nitrogen is present as NH_3 , phosphorus as PH_3 , and germanium as GeH_4 .

The atmospheres of the four planets are reducing atmospheres. Iron oxide, for example, left in such an atmosphere would eventually be reduced to iron, with the oxygen combining with hydrogen to form water. This is in contrast to Earth's atmosphere where, as you know, iron is oxidized to an oxide (rust!)

By studying the spectra of these planetary atmospheres, we can thus obtain the composition of the layers above the clouds, which corresponds roughly to the layers above the 1 bar pressure region. Similar methods are used to obtain the composition of the atmosphere of Titan, Saturn's largest satellite.

7.3.2 The atmosphere of Titan

The structure of Titan is more like that of the terrestrial planets (Figure 7.8). It is believed to have a rocky core of about 1 700 km radius surrounded by a layer of clathrates and/or liquid icy materials. There is a well-defined surface between this layer and the gaseous atmosphere, which has a column mass of $1.09 \times 10^5 \text{ kg m}^{-2}$ (greater than that of the Earth!). Compounds formed in Titan's atmosphere probably precipitate on to the surface as rain or snow to produce solid deposits and/or lakes or seas. TV programme 5 discusses the question of what a probe landing on the surface will find. The probe may splash down into a sea of ethane and methane, hit a solid layer of tar-like hydrocarbons (compounds containing only carbon and hydrogen) or crash on to the 'bedrock' of ice.

As you saw in Subsection 6.5.5, clathrates are solids, such as water ice, with gas molecules, such as N_2 , trapped in them

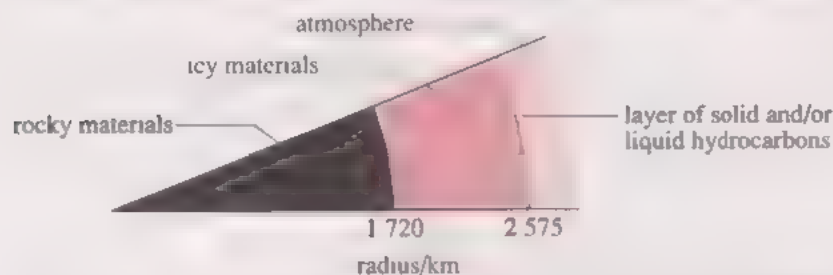


Figure 7.8 The structure of Titan. Titan has a rocky core and an icy mantle surrounding this. The atmosphere lies above the solid icy surface.

The visible layer of Titan's atmosphere (shown in Plate 2 33) is a haze of hydrocarbon droplets. A large number of molecules have been identified in the surrounding gas. These are listed in Table 7.2.

Table 7.2 Composition of Titan's atmosphere

Molecule	Abundance/fraction of total number of molecules
N_2	0.97
CH_4 (methane)	3×10^{-2}
H_2	2×10^{-3}
CO (carbon monoxide)	6×10^{-5}
C_2H_6 (ethane)	2×10^{-5}
C_2H_4 (ethene)	4×10^{-7}
C_2H_2 (ethyne)	2×10^{-6}
C_3H_8 (propane)	$2-4 \times 10^{-6}$
HCN (hydrogen cyanide)	2×10^{-7}
CH_3CCH (propyne)	3×10^{-8}
CHCCCH (butadiyne)	$1-10 \times 10^{-8}$
C_2N_2 (cyanogen)	$1-10 \times 10^{-8}$
HCCCN (cyanoacetylene)	$1-10 \times 10^{-8}$
H_2O	$<1 \times 10^{-9}$
CO_2	$3-7 \times 10^{-10}$

- ☐ Are the molecules in Table 7.2 of elements in oxidized or reduced forms?
- ☒ Mostly in reduced forms.

The main component of Titan's atmosphere is N_2 .

- ☐ Which planet has an atmosphere mainly of N_2 ?
- ☒ The Earth!

The other atmospheric molecules on the Earth, however, are mostly in an oxidized form – CO_2 , O_2 , etc.

□ In what form is nitrogen found in the atmospheres of the giant planets?

■ NH_3 , ammonia.

The discovery that N_2 rather than NH_3 was the dominant form of nitrogen in Titan's atmosphere came as a surprise, and needed an explanation in terms of the chemical processes occurring in the atmosphere, as we shall see in Subsection 7.6.1. The other notable feature of Table 7.2 is the large number of molecules identified, in particular the variety of carbon-containing molecules.

Summary of Section 7.3 and SAQs

- 1 The atmospheres of the giant planets have hydrogen, H_2 , and helium as their major components.
- 2 Other molecules detected are reduced forms of the heavier elements, for example CH_4 and NH_3 .
- 3 The major component of the atmosphere of Titan is N_2 , but other elements are present in reduced forms, such as CH_4 .
- 4 Most of the molecules in the atmospheres are detected by infrared or ultraviolet spectroscopy. Indirect means have to be used for helium atoms.
- 5 Helium is depleted in the atmosphere of Saturn compared with the solar abundance. This may be due to the differentiation of hydrogen and helium in the interior of the planet.

SAQ 7.6 (Objective 7.5) Outline reasons why it might be difficult to detect neon, Ne, in the atmospheres of the giant planets.

SAQ 7.7 (Objectives 7.4 and 7.5) No oxygen-containing molecules have been detected in the atmosphere of Uranus, but this does not mean that there is no oxygen in Uranus. Why not?

SAQ 7.8 (Objective 7.5) Nitrogen molecules in the atmosphere of Titan were detected by the ultraviolet spectrometer on Voyager 1. Why would it be difficult to detect the nitrogen molecules using ground-based instruments on Earth?

7.4 How the atmospheres vary with depth

7.4.1 The giant planets

Although the information on chemical composition refers mainly to the atmosphere above the outermost cloud layer, it has been possible to make measurements of temperature and pressure deeper into the atmospheres. Figure 7.9 shows how the temperature varies with depth above the outermost cloud layers and the layers just below this, for the four giant planets. The curves shown are averages for the whole planet; there are some variations in temperature from pole to equator at any one pressure.

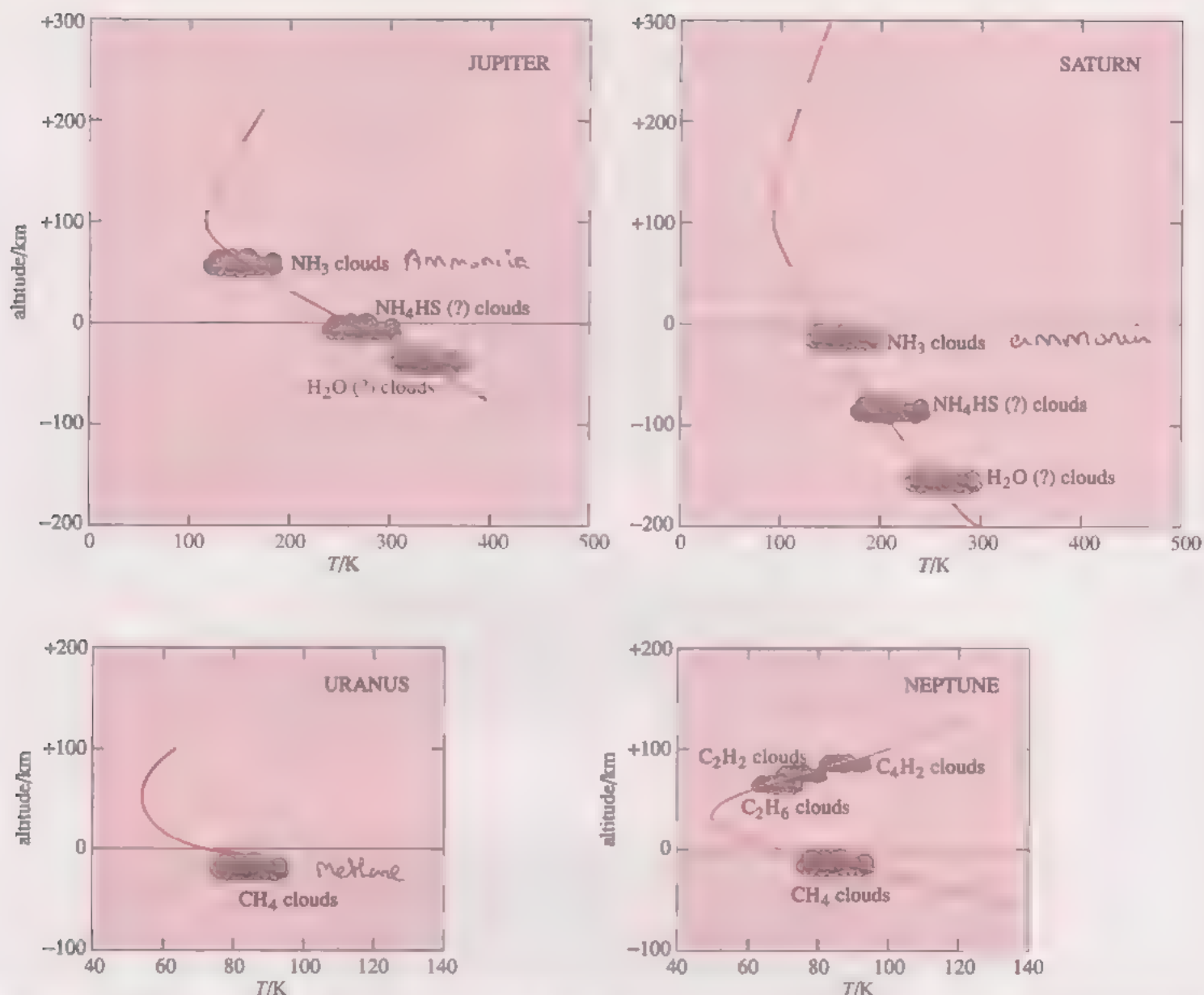


Figure 7.9 Vertical profiles of the atmospheres of Jupiter, Saturn, Uranus and Neptune. Zero altitude is taken to be where the pressure is 1 bar; the profiles extend below this in those cases where measurements are available.

All four curves show a fall in temperature with increasing altitude at the lower depths and an increase in temperature in the outer atmosphere. The atmospheres can thus be divided into a troposphere (Subsection 6.4.2) and, above this, a thermosphere. In the troposphere the decrease in temperature with altitude corresponds approximately to that expected for a gas in which heat is being transported upward by convection.

Much of this information came from radio signals sent through the atmosphere from spacecraft; the amount by which the radio waves are deflected depends on the temperature as well as the molecular mass. Signals sent through at different latitudes pass through different vertical layers (see Figure 7.10) and so give a measure of temperature variation with depth. It was possible to identify the topmost layer of cloud, but the nature of the lower cloud layers marked in Figure 7.9 is based on models of the atmosphere. The radio measurements showed that there are dense cloud layers, which absorbed the radiation, but the substances present were not identified.

The usual way to model the lower cloud layers is as follows. It is assumed that the elements are present in their relative solar abundances. The temperature and pressure are known so we now need to decide in what form the elements will

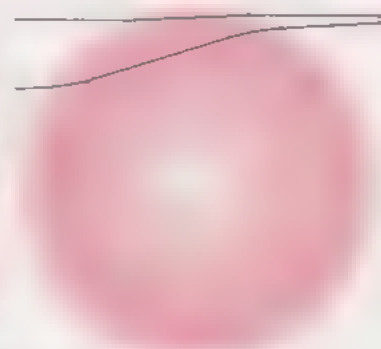


Figure 7.10 Signals at different latitudes pass through different layers of atmosphere. The depths of the layers are exaggerated to show the rays.

be present – atoms or molecules, gas, liquid or solid. The atmospheres are very well mixed and have been present for a very long time, and so they would be expected to have reached chemical equilibrium in the layers beneath the clouds. Chemical equilibrium is explained in Box 7.2.

Box 7.2 Chemical equilibrium

If we take a box containing a mixture of chemicals, shake it well and leave it for a very long time, ensuring that the temperature and pressure remain constant and that no chemicals or radiation leave or enter the box, then the chemicals will reach equilibrium. At equilibrium, chemical reactions will be taking place but the total rate of production of each compound equals the total rate of its destruction so that the amounts of the various compounds present will stay the same. For any chemical reaction occurring, the relative amounts of the compounds involved that are present at equilibrium are given by the **equilibrium constant**, K , for that reaction. The value of K varies with the temperature, but does not depend on the total amount of chemicals present.

Let us take as an example of chemical equilibrium the reaction between nitrogen, N_2 , and hydrogen, H_2 , to form ammonia, NH_3 . We can put into our box any amount of nitrogen, hydrogen and/or ammonia. It does not matter whether we start with a mixture of nitrogen and hydrogen, or with ammonia, or with a mixture of hydrogen and ammonia, so long as both elements are present in sufficient amounts. After a very long time, we will have nitrogen, hydrogen and ammonia present in equilibrium amounts. If we measured the concentrations of the three compounds at equilibrium we could obtain the **equilibrium constant**.

In a planetary atmosphere, there is not just one chemical reaction to consider but a very large number, and any one compound may be involved in more than one equilibrium. Even in the simple case of a mixture of nitrogen and hydrogen, we have to think about compounds other than ammonia that might be formed.

For example, we have to consider the possibility that hydrazine, N_2H_2 , might be formed as well as ammonia. Given the equilibrium constants for the formation of ammonia and for the formation of hydrazine, we can calculate how much of the nitrogen and hydrogen forms ammonia, how much forms hydrazine, and how much is left as N_2 and H_2 .

From our knowledge of how chemicals react, we can choose for any planetary atmosphere a set of reactions involving the most likely molecules formed from the most abundant elements. Even if we disregard some elements, those for example whose abundance is less than a millionth of that of silicon, there are still a large number of reactions, and we need a computer to obtain the **predicted molecular abundances**.

The calculations start where the temperature is relatively high (one very comprehensive study started at 2000 K). The equilibrium constants of all the reactions thought to be likely and the relative abundances of the elements are fed into the computer and the relative molecular abundances calculated. At high temperatures all the icy materials will be gaseous.

We then imagine a parcel of this hot atmosphere rising into the cooler layers. As the temperature falls, the equilibrium constants change and so there is a slight reshuffling amongst the different molecules. A more dramatic change occurs when molecules condense out. Table 7.3 compares the predicted abundances of various molecules in the gas at the 300 K and 150 K levels in Jupiter's atmosphere. The observed abundances are also given for comparison – these are mostly for the atmosphere at cloud-top level (150 K) and farther out, but

in the case of water, ammonia and hydrogen sulphide (H_2S), the observed values refer to levels with pressures of roughly 1–4 bar (150–300 K).

Table 7.3 Predicted abundances in Jupiter's atmosphere

Molecule	Predicted at 300 K	Predicted at 150 K	Observed
H_2	0.89	0.89	0.90
He	0.11	0.11	0.10
CH_4	6×10^{-4}	6×10^{-4}	1.75×10^{-3}
NH_3	1.5×10^{-4}	1.2×10^{-5}	1.8×10^{-4}
H_2O	10^{-3}	8.8×10^{-13}	$1\text{--}30 \times 10^{-6}$
GeH_4	$\leq 6.3 \times 10^{-9}$	~ 0	7×10^{-10}
H_2S	2.9×10^{-5}	8.8×10^{-13}	$< 6 \times 10^{-9}$

Which molecules are predicted to have a very marked drop in abundance on going from 300 K to 150 K?

■ H_2O and H_2S .

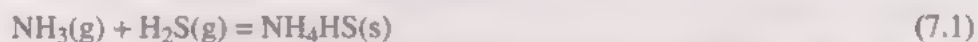
Why do you suppose that these molecules are particularly affected?

■ These are molecules believed to form clouds at layers between 300 K and 150 K (Figure 7.9: the H_2S is incorporated into the NH_4HS cloud).

The NH_3 abundance also drops, but not as dramatically because the NH_3 clouds lie at or near the 150 K level.

The water and ammonium hydrogen sulphide (NH_4HS) layers in Figure 7.9 are predicted from this type of calculation. There is no observational evidence for the composition of these clouds. Consider, for example, the NH_4HS layer.

Ammonium hydrogen sulphide, when pure, is a white solid formed by the reaction of ammonia with hydrogen sulphide:



Equation 7.1, like all chemical reactions, has an equilibrium constant, and this tells us how much hydrogen sulphide must be present if the ammonia is to react and form appreciable amounts of ammonium hydrogen sulphide. The Voyager instruments failed to detect H_2S . Unfortunately, the equilibrium constant for Equation 7.1 leads to a prediction of a detectable concentration of H_2S , where the spacecraft could see through the top layer of clouds.

One possible reason for the apparent lack of H_2S (left over from cloud formation) lies in the colour of Jupiter's clouds. The detected cloud layer is of ammonia, and layers of ammonium hydrogen sulphide and water ice are predicted. All of these substances are colourless. One suggestion is that hydrogen sulphide undergoes chemical reactions that lead to the formation of compounds in which several sulphur atoms are joined together. Some examples are given in Figure 7.11, and the colour of Jupiter's clouds can be seen in Plate 2.34.

The molecules and ions in Figure 7.11 are coloured. If these substances are formed it would explain both the colour of the clouds and the lack of H_2S . There are, however, other possible candidates for the origin of the colour, including phosphorus compounds and carbon compounds, and the spectral evidence is inconclusive.

On Uranus and Neptune only the methane cloud (Figure 7.9), and a methane haze that lies above this cloud, have been positively identified.

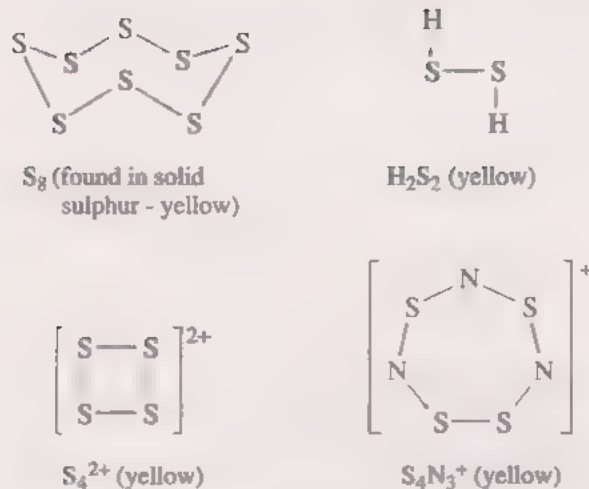


Figure 7.11 Structures of some molecules containing sulphur chains, which are coloured. Molecules such as these may be responsible for the colour of the clouds on Jupiter.

7.4.2 Titan

The variation of the temperature with altitude for Titan is shown in Figure 7.12. These data were obtained from radio wave measurements made by Voyager 1. Unlike the corresponding figures for the giant planets, Figure 7.12 starts at the actual surface of Titan.

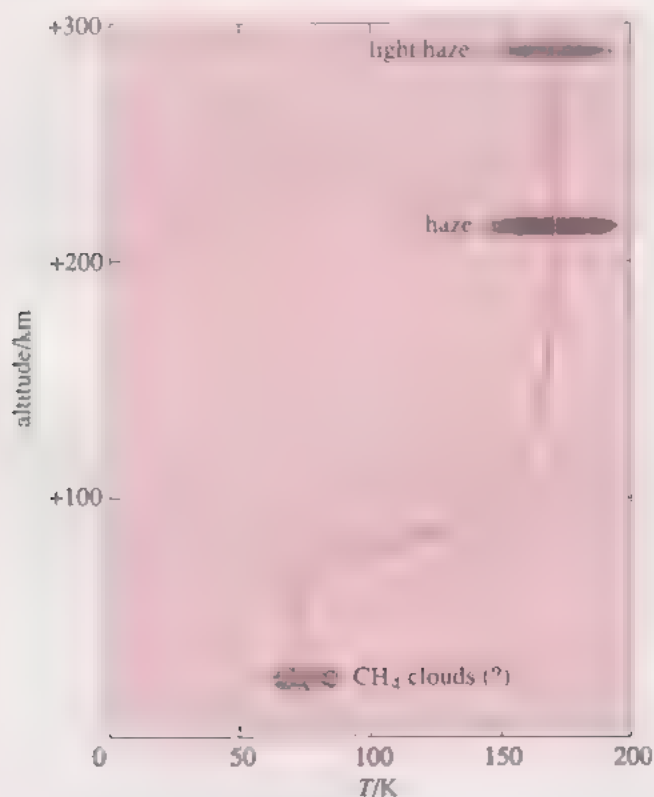


Figure 7.12 The variation of average temperature with altitude in the atmosphere of Titan. Zero altitude is taken at the surface.

11Q 7.4 Add the labels 'troposphere' and 'thermosphere' to the appropriate sections of the curve in Figure 7.12. Does the curve resemble any of those given for the terrestrial planets in Chapter 6 (Figure 6.6)?

There is a haze, believed to be of hydrocarbons, which is very high up. The haze can be seen in the Plate 2.33 as a bluish layer above the yellow clouds. The clouds might consist of methane.

Summary of Section 7.4 and SAQ

- 1 The variation of temperature with depth on all four giant planets divides the atmospheres into two regions. In the lower part, the troposphere, temperature decreases the farther out from the centre we go. The decrease is close to that expected for a region in which convection occurs. In the upper layers of the atmosphere, the thermosphere, the temperature increases with distance from the centre.
- 2 The outermost cloud layers have been identified as ammonia on Jupiter and Saturn and as methane on Uranus and Neptune.
- 3 Models assuming chemical equilibrium can predict the composition of the lower cloud layers, but these compositions have not been positively identified by observation.
- 4 The temperature profile for Titan has been established down to the surface of the satellite. It shows a troposphere and a thermosphere, resembling those in the atmosphere of Mars.

SAQ 7.9 (Objective 7.6) Methane is a relatively abundant molecule in the atmospheres of Jupiter and Saturn. However, no methane cloud layers are predicted there, although such clouds have been detected on Uranus and Neptune. What is the reason for this?

7.5 Atmospheres in motion

When we look at the giant planets through a telescope, what we see is the outermost layer of clouds, but these clouds are not stationary features. They are circulating round the planets, sometimes with speeds in excess of those found for hurricanes on Earth, and the speed changes with latitude. Video sequence 7 shows the motion of the clouds in close up, and you will be asked to watch it (again) shortly. These Voyager images reveal regions of circular motion, called eddies, especially at the boundaries between different cloud bands. The most famous example of such motion is the Great Red Spot of Jupiter, about which we shall have more to say. First, however, we must define what is meant by a wind speed on these planets.

If we wanted to measure the wind speed on Earth from the motion of clouds, we could stand on the surface and measure the speed at which the clouds passed us.

□ Would this be the speed of the cloud observed by a passing spacecraft?

■ No. Standing on the surface we would be moving with the surface so that we would measure the speed of the cloud relative to the surface. The spacecraft measurement would be with respect to a point outside the Earth and so would include the Earth's motion relative to the spacecraft as well as that of the cloud with respect to the surface.

Now, the giant planets probably do not have surfaces, and certainly no surface has been detected. So how are we going to allow for the rotation of the planet?

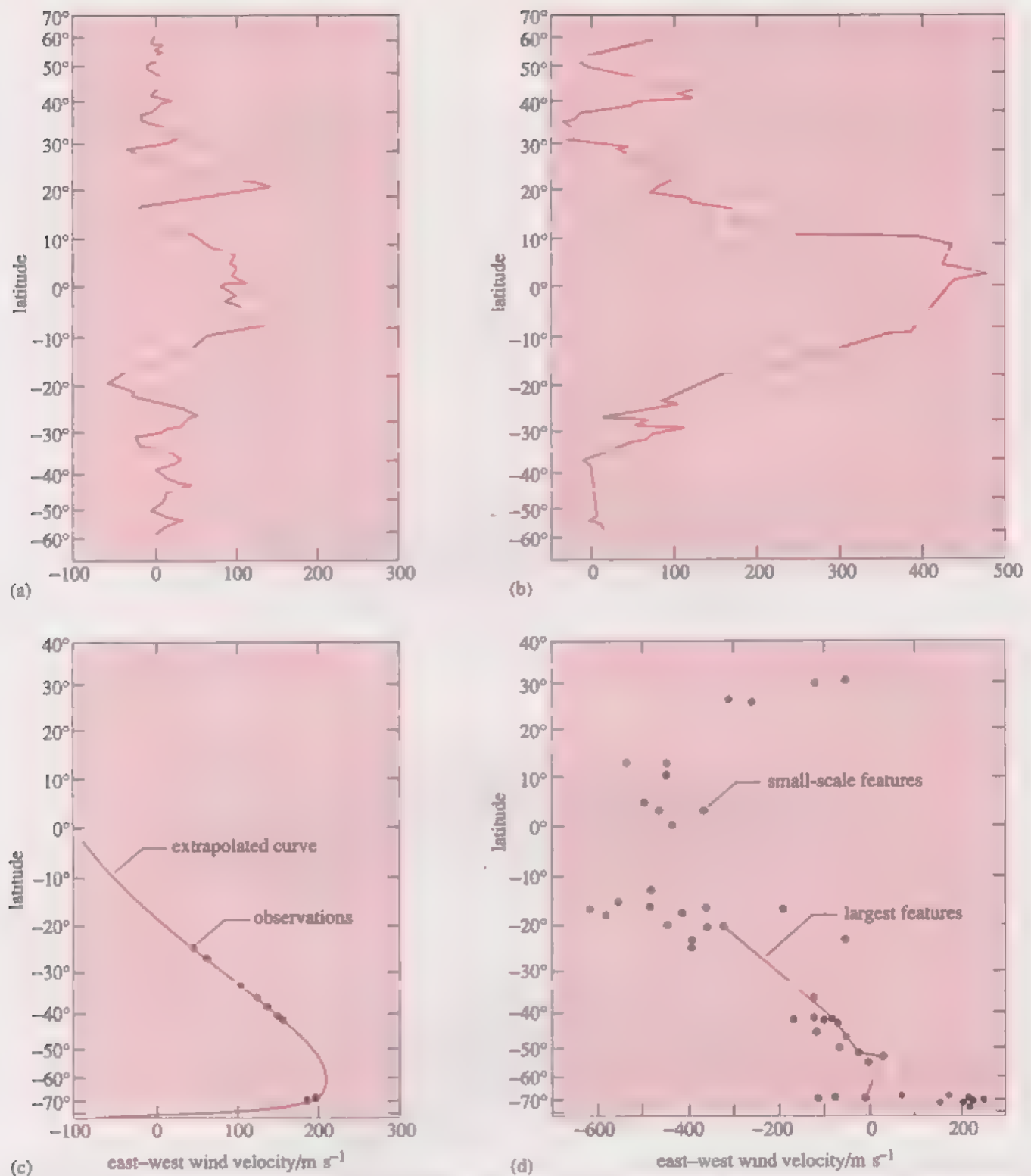


Figure 7.13 Measured east-west wind velocities at different latitudes for (a) Jupiter, (b) Saturn, (c) Uranus and (d) Neptune.

Well, what is done is that the rotation of the interior of the planet is measured and then subtracted from the speed of the clouds to give the wind speed.

An ingenious method has been employed to measure the rotation of the planetary interior. Around the planets are charged particles, electrons and ions. These come from various sources, such as the solar wind, cosmic rays, the planet itself and (in the case of Jupiter) volcanic activity on Io. These particles are accelerated by the magnetic field of the planet and, as a consequence, emit

radiation. As the planet rotates, the magnetic field rotates with it. Because the field is inclined to the spin axis, the electrons and ions experience a magnetic field that varies with the planet's rotation. This leads to bursts of radiation whose interval gives the rotation speed. The period of rotation of all four planets has been measured in this way by monitoring bursts of radio frequency radiation, notably by spacecraft. The periods of rotation were given in Table 1.1a. Note that, despite their greater size, all four planets have shorter rotation periods than the Earth.

The measured wind velocities from equator to pole are shown in Figure 7.13 for the four planets. In Plate 2.36 these are shown superimposed on images of Jupiter and Saturn so that you can see how changes in wind velocity match changes in cloud colour or other features.

To help you to understand wind speed and to gain a general impression of atmospheric motion on the giant planets, you should now watch video sequence 7, Atmospheric circulation, and do the exercise.

7.5.1 Jupiter

The visible layer of Jupiter shows a pattern of alternate dark and light bands parallel to the equator called **belts** and **zones**, respectively. It has been shown that in the light zones the atmosphere is rising and in the dark belts it is sinking. Plate 2.36 shows that changes in east–west wind velocity often occur at the boundaries between belts and zones. These observations lead to a picture of a series of convection cells (Figure 7.14). Convection in the form of Hadley cells, as you may recall from Subsection 6.4.4, is a feature of atmospheric motion on the terrestrial planets.

As the atmosphere rises it expands and cools; ammonia condenses out and forms white clouds, which give the zones their light appearance. The atmosphere cools radiatively and begins to sink. When it sinks in the belts the atmosphere is compressed and so gets hotter, and therefore the clouds evaporate so that we can see the coloured material underneath. What is driving these convection cells and why are there so many?

On Earth, convection cells are formed when air is heated at the Earth's surface. What sources of heating are available on Jupiter?

■ There is no surface, but some of the cloud layers might absorb solar radiation (and emit infrared). The internal heat of the planet will contribute [Comment: There is also some heat given out when gases condense to form clouds.]

The most significant source of heating is the internal heat of the planet. Convection would be the most efficient way of transporting this heat from the interior to the outer atmosphere. The vertical temperature profile of the troposphere shows that the rate of cooling across this region is that expected for a convection process, so we can be fairly sure that convection cells occur here. It is possible that the cells extend deep into the interior, but we have no measurements of motion below the troposphere.

The Coriolis effect (Subsection 6.4.4) plays a major role on the giant planets. Its role on Earth arises from the rapid rotation of the planet.

ITQ 7.5 Jupiter has a radius (at the 1 bar level) of 71 400 km at the equator. Its period of rotation is 0.410 days. At what velocity (with respect to a non-rotating observer in space, moving at the same speed as Jupiter in its orbit) would a still atmosphere at the equator be moving?

The term **velocity** is used when we want to indicate the direction of motion as well as its magnitude (speed). In Figure 7.13 we distinguish winds in the same direction as the rotation of the planet (positive velocity) and those in the opposite direction (negative velocity).

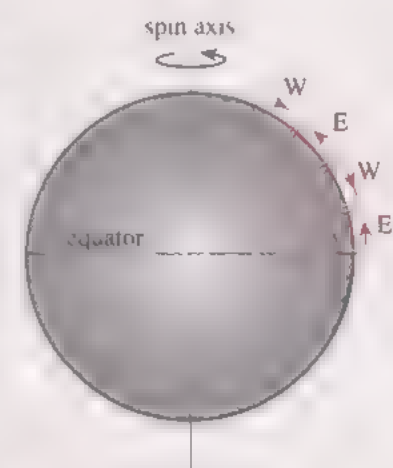


Figure 7.14 Schematic convection cells for Jupiter. Note that the high rotation speed of Jupiter leads to a succession of small cells. E means that there is flow to the east, and W to the west.

The equivalent figure for Earth is 465 m s^{-1} . The Coriolis effect is thus much greater for Jupiter than for the Earth. Computer simulation of the atmosphere based on convection cells formed in only the outer layers of Jupiter shows that, at the rotation speed of Jupiter's atmosphere, the Coriolis effect no longer leads to a break-up of the cells as occurs on the Earth north and south of the tropics, but to a banded structure of small convection cells. This is shown in video sequence 7. This is not the only model that can reproduce the banded structure, however. One alternative considers Jupiter as a spinning, fluid sphere. Convection in the interior of such a sphere occurs in a series of co-axial cylinders about the spin axis. Each cylinder reaches the surface at a different latitude and forms a zone, as shown in Figure 7.15.

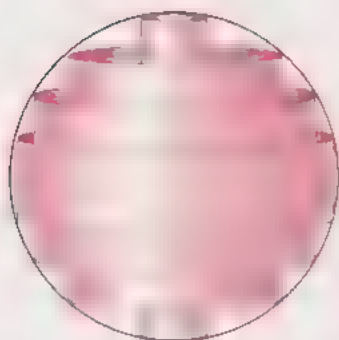


Figure 7.15 Co-axial cylinders in a fluid sphere.

The beautiful close-up images from Voyager (Plate 2.35) enable us to see that there are also regions of turbulent motion, particularly at boundaries of zones and belts. These are storms resembling hurricanes on Earth. On Earth, such storms die out fairly quickly as they lose energy to the surface. There are no surface features at the base of the storms on Jupiter and so they do not lose energy through surface friction. In addition, the lower temperature of Jupiter's atmosphere means that less energy is lost by being radiated out into space. The energy from the storms appears to feed back into the zones to sustain the high wind speeds. The most famous storm is the Great Red Spot, which we consider in more detail now.

7.5.2 The Great Red Spot

The **Great Red Spot** is a prominent feature of the Jovian atmosphere. It covers a vast area, being some 14 000 km from north to south and 26 000 km from east to west. (In either direction, it spans a distance greater than the diameter of any terrestrial planet.) Because of its size and the contrast in colour between itself and the surrounding atmosphere, it can be observed by ground based telescopes of only moderate power. The first definite record of the Great Red Spot dates from 1830, but the earliest sighting of a large spot on Jupiter is thought to be by the Italian-French astronomer Jean Cassini (1625–1712) in 1665. Thus the Great Red Spot has been present for 160 years and possibly for more than 300 years. This would not be surprising if it were a surface feature, such as the mountains of Earth or Mars, but the close-up of the Great Red Spot (Plate 2.40), and the time-lapse pictures in video sequence 7 show clearly that it is a region of circulating atmosphere, a giant storm. It is rotating anticlockwise and completes one rotation every six days.

As you saw in Subsection 7.5.1, storms on Jupiter dissipate their energy slowly because there is no surface, and because the low temperatures mean that the atmosphere is radiating less heat to space. Nonetheless, it is believed that the Great Red Spot must be shallow because the pressure increases rapidly through the atmosphere of Jupiter and would soon be great enough to disrupt the circulation. The drift of the Spot in longitude shown by historical and modern records also indicates that the storm cannot be deep-seated. It is possible, however, that there is a large deep convective cell below the Great Red Spot, which causes that part of the atmosphere to be a 'hot spot' and provided the energy for the initiation of the Spot.

Computer modelling has shown that large, stable eddies can be produced regardless of the process maintaining the winds. Thus the presence of the Great Red Spot does not allow us to decide between competing models of Jupiter's atmosphere. More detailed analysis of how the Great Red Spot interacts with the surrounding atmosphere may enable us to narrow down the possible models.

There remain many unsolved problems. For example, why is there only one Great Red Spot, why not one in each hemisphere? The colour of the spot is also puzzling. The temperature of the Spot as measured by infrared spectrometers is the same as that of the surrounding white ovals, which are also storms, indicating

that these features are all at the same depth. Yet on Jupiter as a whole the coloured layers are deeper and at a higher temperature than the white clouds. Is the colour due to the same material as in the coloured cloud layers? If so, why does it occur at a different depth in the Great Red Spot?

We shall have to leave you with these questions unanswered and turn now to the other giant planets. How do their winds compare with Jupiter's? Does any other giant planet have a feature comparable with the Great Red Spot?

7.5.3 Winds on Saturn

Saturn is in many ways similar to Jupiter, and a similar pattern of winds might therefore be expected. To some extent, these expectations are fulfilled. Saturn's atmosphere is banded, and contains streams moving at different velocities. The pattern is notably different from that on Jupiter, however. A major difference is the stream of very high speed wind near the equator. As you can see in Figure 7.13, the wind speed reaches 500 m s^{-1} (which is about two-thirds of the speed of sound on Saturn). The highest wind speeds on Jupiter are about 150 m s^{-1} . None of the models put forward to explain the banding in the atmospheres of Jupiter and Saturn can account for this.

The other main difference is that, unlike on Jupiter, the changes in wind velocity on Saturn are not closely allied to zone boundaries. The contrast between belts and zones is much more muted. A true-colour image shows a fuzzy, orangey planet with only very faint markings (Plate 2.34). If we image the planet at several different wavelengths, and then recombine the different images with different intensities, the banded structure becomes apparent as in Plate 2.35. As Plate 2.36 shows, however, the different colour stripes do not correlate particularly well with different wind velocities.

Eddies appear on Saturn, as for example in Plate 2.41, where exaggerated colours have been used to show up the features. The largest found was 6 000 km long, considerably smaller than the Great Red Spot.

In 1990, a Great White Spot appeared on Saturn and over the course of about a month spread to encircle the whole planet (Figure 7.16). Plate 2.42 shows Saturn in the aftermath of this storm. A similar feature had been observed earlier at roughly 30 year intervals (1933, 1960) but was not present when either Voyager probe flew by (December 1980 and August 1981). (The 1933 sighting, incidentally, was first reported by an amateur astronomer, the Scottish comedian Will Hay.) The Spot forms at the north boundary of the equatorial zone, thus reinforcing the theory that such features form when two streams of different velocity meet. The atmosphere is rising in this region, and the white colour is due to ammonia crystallizing out.

7.5.4 The atmospheric circulation of Neptune

Rather than continue outwards through the Solar System to Uranus, we jump to Neptune, leaving Uranus until later. This is because the data on atmospheric motion on Uranus are extremely sparse owing to the lack of discernible features. Neptune, like Jupiter and Saturn, has a banded structure, in this case patterned in blue and white (Plate 2.39). By using colour-enhanced images it has been possible to discern several cloud features. Compare, for example, images of Neptune through an orange filter and through a green filter (Plate 2.38).

Why is it important to be able to distinguish features such as spots in the cloud layers?

■ The velocities of the cloud layers are measured by tracking features across the planet as it rotates. If the cloud layers appear uniform then we cannot track them.



Figure 7.16 The Great White Spot on Saturn. These images from the European Southern Observatory show the rapid expansion of the Great White Spot during October 1990.

Unfortunately, many of the features at the time of the single Voyager fly-by (August 1989) were relatively short-lived, and wind velocities obtained from small-scale, short-lived features and the longer-lived large-scale features do not always coincide. This may be due to the small-scale features moving relative to the wind. There is sufficient information to draw some conclusions, however.

□ Look carefully at the wind velocity scales in Figure 7.13, and state one obvious difference between the winds on Neptune and those on Jupiter and Saturn.

■ The wind velocities on Neptune, except near the poles, are negative. That is, the atmosphere is rotating more slowly than the interior.

On Jupiter and Saturn there are both positive and negative wind velocities, but the majority are positive. The equatorial wind velocity on Neptune, as well as being negative, is also very high, at least 300 m s^{-1} and possibly 560 m s^{-1} . It seems then that Neptune has an equatorial stream like Saturn, but in the opposite direction.

The negative equatorial velocities are more easily understood than Saturn's high positive equatorial velocities, because they can be explained in terms of convection in the atmosphere. When the measurements were taken the Sun was overhead at mid-latitudes in the southern hemisphere, and so these latitudes were being heated more than others by absorption of the Sun's radiation in layers at the base of the atmosphere. The heated atmosphere would then rise at mid-latitudes and travel towards the equator and pole. At the equator, the atmosphere would fall (see Figure 7.17). The Coriolis effect causes the gas to slow down as it travels equatorwards and so it has a negative velocity. Thus at the equator there is a negative wind velocity.

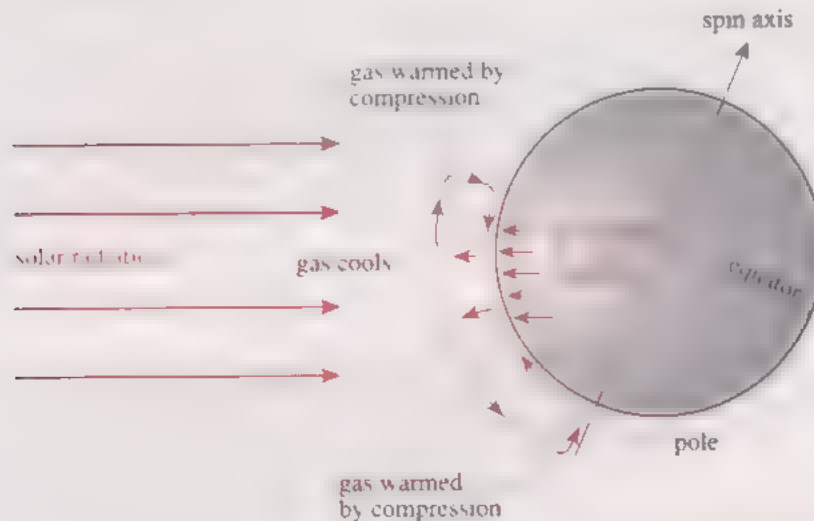


Figure 7.17 Absorption of solar radiation and subsequent atmospheric convection on Neptune.

What would you predict the wind direction to be at the south pole of Neptune, using this explanation?

■ The warmed mid-latitude gas moves polewards. However, this time the gas is moving to a latitude of smaller radius and so will move faster at the pole than at mid-latitudes owing to the Coriolis effect, resulting in positive velocities.

There is some evidence for such a convection process from temperature measurements. The solar radiation will be absorbed at cloud level, where we would expect to find that mid-latitudes were the warmest. As the warm atmosphere at mid-latitudes rises, it cools, so that the temperature at higher altitudes will be lower. When this cooler atmosphere starts to fall at the pole and the equator, it at once starts to be compressed, and this compression causes a rise

in temperature. At the higher altitudes *towards the top of the convection cell*, therefore, the pole and the equator will be warmer than mid-latitudes. The layers that infrared measurements sampled were at higher altitudes than the clouds, in fact where we might expect this compressional heating to take place. These measurements did indeed show that the pole and the equator were warmer (by a few kelvins) than mid-latitudes.

7.5.5 Neptune's Dark Spot

Neptune's Dark Spot is a huge storm in Neptune's atmosphere (Plate 2 39). At first glance it is the local equivalent of Jupiter's Great Red Spot. Figure 7.18 shows an article from the October 1990 issue of the journal *Sky and Telescope*. You are asked to read it and then answer some questions about the Dark Spot. We suggest you first read the article through completely. Then read the questions in the Exercise. For each question go back to the article and try to pinpoint those parts relevant to the question. Make notes, copy out the relevant parts or highlight them. Then decide what you are being told. Use this material and information from earlier in this Chapter to answer the questions in your own words. Examples of this process are given in our answers (p.202).

Jupiter's Red Spot And Neptune's Dark Spot

The realization that Jupiter's Great Red Spot is a storm system so huge that its smallest horizontal dimension is larger than the Earth is enough to pique curiosity. But when the Voyagers flew past Saturn in 1980 and 1981, and then Uranus in 1986, we became more and more aware that the sheer size of the Red Spot made it a unique atmospheric feature. Only three minor, low-contrast, oval weather systems were found in Saturn's less-turbulent atmosphere, and Voyager 2 found Uranus' calm, blue outer layers nearly featureless.

These new observations supported the idea that large weather systems like the Red Spot must form when small eddies are forced together by a tumultuous atmosphere to form larger ones — medium-size eddies then eat little eddies, and big eddies eat medium-size ones. But why should Jupiter, with its multitude of storms, develop only one large red spot? Delighted with our new mother lode of data, and hopeful that we could answer this question, we retreated to our offices, laboratories, and computer facilities to organize our attack.

But we were once again jolted to attention by Voyager's photographic bounty, this time from the August, 1989, encounter with Neptune. The discovery of a Great Dark Spot on Neptune reminded us that much remains to be learned about the nature of the atmospheric circulation of the giant planets. Here, on a planet receiving only about 3 percent of the sunlight received by Jupi-

ter and with only 5 percent of Jupiter's mass, we find the Great Dark Spot — a huge storm rivaled only by the Great Red Spot.

Why should two large, oval weather systems appear at about the same latitude on the two planets (about 20° S) and cover similar portions of the surface area of the cloud deck? And why just on Jupiter and Neptune — why not on Saturn and Uranus? Are the two giant spots similar, or are differences in atmospheric transparency and structure forcing us to sample these storms at such different levels that comparison of their fundamental natures will be difficult?

Actually, while the giant Neptunian storm may lie at nearly the same planetary latitude as Jupiter's Red Spot, it is dissimilar in several ways. The Dark Spot is not as tightly constrained by the surrounding winds. Time-lapse movies assembled from Voyager images reveal that the Dark Spot rotates counterclockwise and wobbles as it spins around, instead of keeping a precisely tailored perimeter like the Red Spot (see pages 360 and 364). The Neptunian storm's attendant white clouds appear to form as air rises and flows over and around the storm center without being consumed by it. These cirruslike wisps are highly variable, condensing or evaporating in only an hour or so, and serve as poor markers with which to determine the Dark Spot's circulation. There is little visible evidence of strong wind jets to the north and south of the spot, even though it

must be nested in such a pattern. The storm's contrast with its surroundings is more like a stain; little structure accompanies the color change. Finally, unlike the Red Spot, the Dark Spot appears bright in ultraviolet light and is darkest in green and blue light.

We must also include the storms of Saturn in this comparative study. Voyager's snapshots revealed an oval cloud system at 50° south latitude spanning about 3,000 kilometers. Although the lack of discrete structures within this cloud system prevents us from determining its sense of rotation, the surrounding wind pattern indicates that it also rolls counterclockwise. Saturn's feature also displays the same well-defined oval perimeter as the Red Spot's, but unlike its Jovian cousin, it does not strongly absorb blue and violet light.

The relative rarity of such features leaves us with the impression that creating large storm systems isn't easy. However, once one forms, it does not vent its rage in a brief tantrum like a terrestrial hurricane. There is a dearth of detail on the internal motions within Neptune's or Saturn's storms. But this is not the case with Jupiter. The fine detail in the Voyager imagery, a long-term record of low-resolution, ground-based observations, the expectation of data from the Hubble Space Telescope spanning more than a Jovian year, and the anticipated yield from the Galileo mission promise that the cloud features may soon yield to our investigations.

Figure 7 18 An article about Jupiter's Great Red Spot and Neptune's Dark Spot, from *Sky and Telescope*

Exercise

Question 1 Are there any features comparable with the Great Red Spot and the Dark Spot on Saturn or Uranus?

Question 2 Why were the investigators surprised to detect the Dark Spot in Neptune's atmosphere?

Question 3 Describe how the circulation in the neighbourhood of the Dark Spot differs from that of the Great Red Spot.

Question 4 What evidence in the article suggests that the coloured material in the Dark Spot is of a different chemical composition from that of the Great Red Spot?

7.5.6 Winds in the atmosphere of Uranus

Find the tilt of the spin axis of Uranus from Table 1.1a. How might this unusual tilt affect atmospheric circulation?

■ The planet is effectively tipped over so that for large fractions of the Uranian year, a pole receives more solar radiation than the equator. Figure 7.19 shows the inclination at the time of the Voyager 2 observations (January 1986). This could lead, for example, to a convection cell originating at the pole and producing negative wind velocities at mid-latitudes through the action of the Coriolis effect.

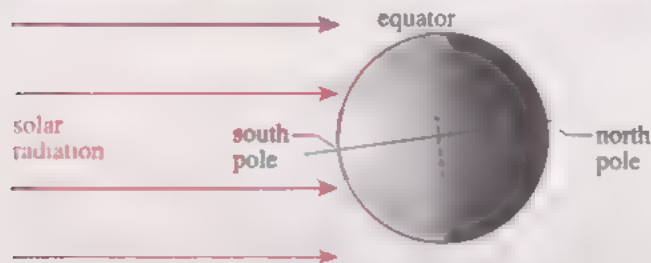


Figure 7.19 The tilt of Uranus, showing how one pole receives more solar radiation than the equator at certain times during the planet's year. Uranus was tilted this way when Voyager 2 passed by.

What other features of Uranus might cause the pattern of atmospheric circulation to differ from that on the other three giant planets?

■ Uranus has a negligible heat excess. The internal heat source may be a significant factor in atmospheric circulation.

To a passing astronaut, Uranus would appear to be enveloped in a blue-green fog in which virtually no features are discernible, as in Plate 2.34. In Plate 2.37 it is possible to make out a banded structure similar to those on the other three planets. This picture was obtained by combining three images of Uranus: one taken with a violet filter, one with an orange filter and one with a filter for radiation of wavelength 619 nm (corresponding to a spectral line of methane). Contrast-enhanced images were recombined, with the violet filter image shown as blue, the orange filter image as green and the methane line image as red. Such methods made it possible to find a few features, such as small spots.

Figure 7.13c shows the few wind velocity measurements available. These are all positive but do not extend into the equatorial region. Extrapolation leads to a possibility of negative velocities within 20° of the equator. As on Neptune, infrared measurements show that the pole and equator are warmer than mid-latitudes.

] Is this the expected temperature distribution?

■ Not if the explanation for the temperature distribution on Neptune is correct. This depended on mid-latitudes receiving most solar radiation. When Voyager flew past Uranus, the south pole was pointing towards the Sun, so it is surprising that the equator was warmer than mid-latitudes.

Overall, the data we have suggest that the circulation on Uranus is very similar to that on Neptune. The lack of any marked difference between the two remains one of the surprising results of the Voyager missions.

Summary of Section 7.5 and SAQs

- 1 Wind velocities on the giant planets are measured by tracking the movement of cloud features.
- 2 These measurements will give a velocity that includes the rotation speed of the planetary interior and so this has to be subtracted.
- 3 The rotation speed of the interior can be measured from radio bursts.
- 4 All four planets have a banded cloud structure.
- 5 Jupiter and Saturn have positive wind velocities at the equator. Neptune has a negative equatorial wind velocity, and extrapolation suggests that Uranus has a negative equatorial wind velocity too.
- 6 On Jupiter and Saturn, there is evidence for a series of convection cells giving rise to the observed pattern of wind velocities.
- 7 On Jupiter, major changes in wind velocity correlate with the boundaries between different coloured bands. There is no such correlation on Saturn.
- 8 Equatorial wind velocities can be very high, 500 m s^{-1} (positive) on Saturn, and possibly 560 m s^{-1} (negative) on Neptune.
- 9 Jupiter, Saturn and Neptune have large features called Spots, which are huge storms.. Jupiter's Great Red Spot has been observed continually for over 100 years. Saturn's Great White Spot appears only periodically.
- 10 Despite the unusual tilt of the spin axis and the apparent lack of internal heat, there is no evidence for a different circulation pattern on Uranus from that on the other three planets.

SAQ 7.10 (Objective 7.8) What are wind velocities measured relative to for the four giant planets? Does this definition of wind velocity assume anything about the planetary interiors?

SAQ 7.11 (Objective 7.8) Some of the wind velocities in Figure 7.13 are shown as negative and yet all the atmosphere is rotating in the same direction. Explain this.

SAQ 7.12 (Objective 7.9) List some observations on heat sources and heat transport on Jupiter that must be taken into account when developing theories to explain the pattern of winds.

SAQ 7.13 (Objectives 7.8 and 7.9) Explain why it is possible for a hurricane-like storm to exist for 300 years on Jupiter whereas similar storms on Earth die out much more rapidly.

SAQ 7.14 (Objectives 7.8 and 7.9) Would an internal heat source warm the poles, equator or mid-latitudes most? Does the explanation for the high negative wind velocity at the equator on Neptune given above suggest a dominant role in atmospheric circulation of the internal heat source?

7.6 Atmospheric chemistry

In the preceding Section, you have seen that the atmospheres of the giant planets are in constant turmoil. Recorded wind speeds can be very high, and there are vast storms present. In chemical terms, however, it is safe to assume that below the cloud layers, the various atoms and molecules are in chemical equilibrium. Indeed, the very turbulence of the atmosphere aids the reaching of equilibrium by ensuring that the gases are well mixed.

Above the clouds we have to allow for an extra participant in the chemistry that is not available lower down. Can you think what this is?

■ Sunlight.

Although in the dark outer reaches of the Solar System much less sunlight is received than on Earth, the absorption of light still plays an important role in the chemistry. The chemistry of the giant planets' atmospheres is rich and complex and, rather than try to cover all the important reactions, we shall consider two problems that arise from the observed molecular abundances. These are (i) that nitrogen is found mostly as the nitrogen molecule, N_2 , on Titan but as ammonia, NH_3 , on the giant planets, and (ii) that carbon-containing molecules that would not be expected in any measurable amount at chemical equilibrium are found to be present in the atmospheres of Jupiter and Titan in particular. The existence of these latter molecules has led to speculation about the possibility of biologically important molecules being present.

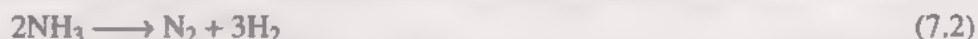
7.6.1 Nitrogen on Titan

Titan is very unusual in having an atmosphere whose main constituent is nitrogen as N_2 . Titan's size and rocky interior resemble the terrestrial planets more than the giants, but the types of molecule (apart from N_2) found in the atmosphere cause the atmospheric chemistry to resemble that on Jupiter rather than that on the Earth.

There are two explanations for the nitrogen atmosphere of Titan. The first is that when Titan formed, the very low temperature of that part of the solar nebula caused the nitrogen to be trapped as a clathrate (Subsection 6.5.5) in the icy layers as the satellite formed.

Now, how does this trapped nitrogen reach the atmosphere? In a clathrate, one substance crystallizes in a structure that contains a lot of spaces or cages. Other substances can be trapped in these cages. Water, as ice, forms many structures of this type. The nitrogen cannot easily leave the ice because there are no tunnels of sufficient size leading out of the cage, but when the ice is warmed and begins to melt, then the nitrogen is lost very readily. On this theory, then, slight warming of the icy layer, for example by radiogenic or tidal heating, could release the nitrogen. Other mechanisms also exist for releasing the nitrogen. On this view, the nitrogen is present as N_2 because it was trapped as N_2 when the satellite was formed. Any ammonia present remained as a solid ice and was never released into the atmosphere.

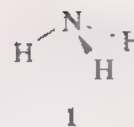
A second model assumes that initially the nitrogen was present as ammonia, and that subsequent reactions converted this into nitrogen. The overall reaction is



but this reaction would be very slow at the temperature of Titan and, if the atmosphere were at chemical equilibrium, there would be substantial amounts of ammonia present. That there are not is due to several factors, including the interaction of sunlight with the atmosphere and the escape of hydrogen. Let's see what happens when sunlight strikes a molecule of ammonia.

Ammonia is a pyramid-shaped molecule with three hydrogens attached to the nitrogen atom (structure 1). If an ammonia molecule absorbs an ultraviolet photon of sufficient energy, it can undergo photodissociation (Subsection 6.5.1).

One way in which photodissociation can occur is illustrated in Figure 7.20. Here, the energy absorbed is channelled into one bond causing it to break.

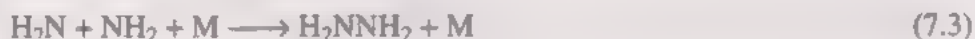


In representations of molecules such as structures 1 and 2, a wedge shape is used to show a bond that is coming out of the plane of the paper, and dashes to show a bond that goes behind the plane of the paper.



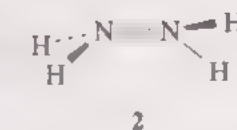
Figure 7.20 Schematic representation of the photodissociation of an ammonia molecule.

The initial products of photodissociation are hydrogen atoms, which escape from Titan, and the radical NH_2 , shown in Figure 7.20. This radical rapidly reacts further, for example with another NH_2 radical to form the molecule hydrazine, N_2H_4 (structure 2), as in Equation 7.3:



□ What is M in Equation 7.3 and why is it there?

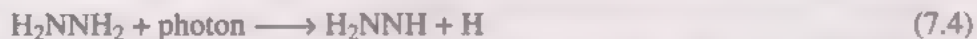
■ M is a third body, a molecule that does not take part in the reaction but is needed to take away some of the energy of the reacting molecules. In the absence of M, hydrazine would dissociate to two NH_2 radicals again as soon as it had formed.



Generally when we write chemical formulae, we group together all like atoms as in N_2H_4 , but sometimes it can make the structure of the molecule clearer if we write out the formula in its constituent parts. Writing N_2H_4 as H_2NNH_2 tells us, for example, that hydrazine consists of two NH_2 fragments joined by a bond between the two nitrogens. This sort of representation is particularly used for compounds of carbon.

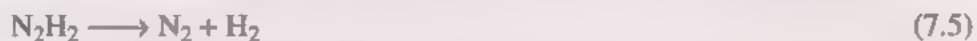
Similar photochemistry for ammonia occurs in the present atmosphere of Jupiter. However, there are two important differences between Jupiter's photochemistry and the proposed route to N_2 on Titan. First, at the temperatures and pressures in the layers of Jupiter's atmosphere where photodissociation occurs, hydrazine condenses out to form a haze. This haze does not react further. It has been proposed that on Titan, by contrast, the hydrazine remained in the gas phase. For this to happen it would be necessary for the surface temperature of Titan to be some 50 K higher than it is at present. Hence the model must include some mechanism for raising the surface temperature.

Secondly, hydrogen as atoms, H, or molecules, H_2 , is more readily lost from Titan than from Jupiter owing to the low escape speed of Titan. On Jupiter, hydrogen is so abundant that the NH_2 radical recombines readily with hydrogen, to re-form ammonia. On Titan, the gaseous hydrazine will itself undergo photodissociation with the loss of a hydrogen atom to form the radical H_2NNH :

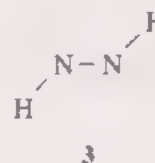


This radical undergoes further reactions with the radical N_2H_3 (i.e. H_2NNH), and with other radicals such as NH_2 or H.

These reactions form N_2H_2 (HNNH , structure 3), this breaks down to give nitrogen and hydrogen:

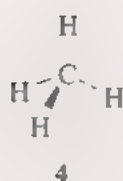


The hydrogen escapes into space, leaving nitrogen in the atmosphere.



7.6.2 Photochemistry of hydrocarbons

As you saw in Table 7.2, the atmosphere of Titan contains a wide range of carbon-containing molecules, in particular **hydrocarbons** (compounds whose molecules contain only carbon and hydrogen atoms linked together). At the temperatures found in Titan's atmosphere some of these condense. Tiny droplets of liquid hydrocarbons form the haze seen in images of Titan, such as Plate 2.33. Many of these compounds will not remain in the atmosphere but fall on the surface as rain or snow. Thus the surface of Titan could be covered by deposits of solid hydrocarbons and lakes of liquid hydrocarbons. The solids could be crystalline deposits of hydrocarbons with only relatively few carbons in each molecule, rather like snow, or tarry deposits of hydrocarbons containing long chains of carbon atoms.



If, however, the atmosphere were in chemical equilibrium, the only hydrocarbon present would be methane, CH_4 (structure 4), the simplest hydrocarbon. So why are the other hydrocarbons present? These molecules are the result of photochemical reactions in the outer atmosphere. In those regions where solar radiation (particularly ultraviolet) penetrates, atmospheric composition is determined not by chemical equilibrium but by the interaction of molecules with radiation. For example, assuming chemical equilibrium, the predicted fractional abundances of ethane, C_2H_6 , and ethyne, C_2H_2 , are negligible, but if the effect of ultraviolet radiation is included, then the predicted fractional abundances rise to 10^{-5} and between 10^{-8} and 10^{-6} , respectively. These figures agree reasonably well with the observed fractional abundances in Table 7.2. How, then, can the observed variety of hydrocarbons be produced through the action of solar radiation on methane?

Let's start by seeing what happens when methane is photodissociated. The methane molecule has a central carbon atom joined to four hydrogen atoms, which can be thought of as lying at the corners of a tetrahedron (structure 4).

Now methane can absorb ultraviolet radiation and break up into smaller fragments in several ways. We can imagine, for example, one bond acquiring so much vibrational energy that the molecule cannot contain it and a hydrogen atom separates off. However, the C—H bonds do not necessarily vibrate in isolation, and we can also have processes in which two, three or four bonds vibrate together. The most common photodissociations of methane in the atmospheres of Titan and the giant planets are those in which two hydrogens are lost. These are illustrated in Figure 7.21, and in Equations 7.6 and 7.7.

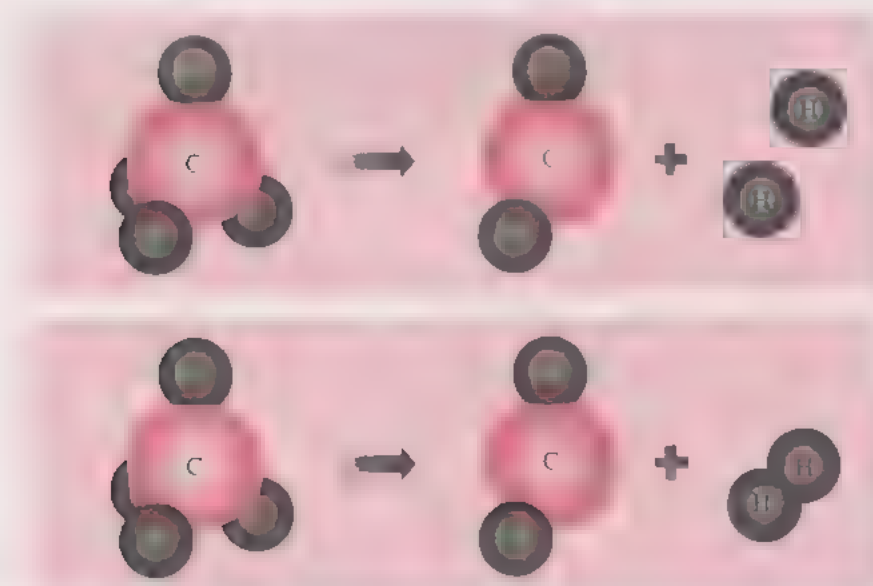
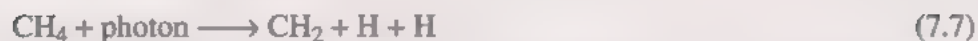
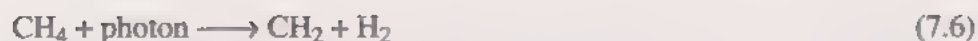
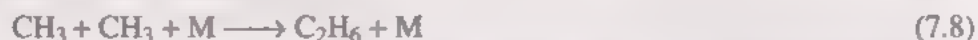


Figure 7.21 Schematic representation of the photodissociation of methane.



In the atmospheres of the giant planets the most abundant molecule is H_2 . This reacts with the carbon-containing product of Equations 7.6 and 7.7, the methylene radical CH_2 . The two main products are methane and the CH_3 (methyl) radical. Two methyl radicals will combine to form ethane, C_2H_6 (structure 5), as in Equation 7.8:



The formation of methane, of course, just takes us back to the starting material.

A more interesting set of reactions occurs when methane loses three hydrogens to form CH (Figure 7.22). This happens about 8% of the time when methane absorbs ultraviolet radiation, and leads ultimately to the production of hydrocarbons containing long chains of carbon atoms.

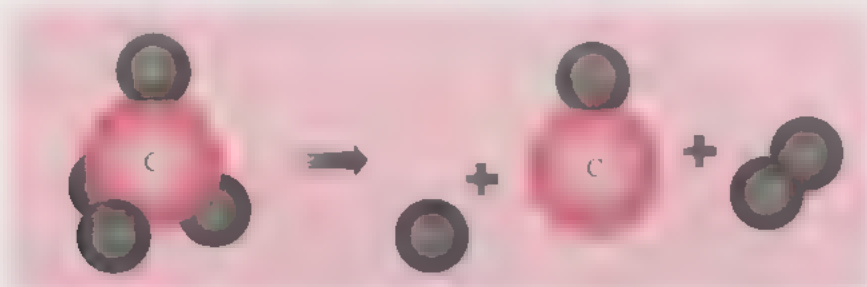
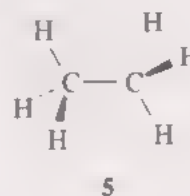


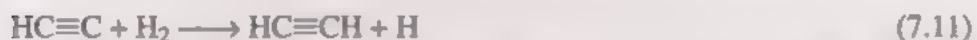
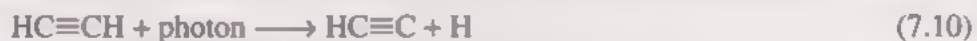
Figure 7.22 Schematic representation of the photodissociation of methane to form CH .

The CH produced by loss of three hydrogen atoms reacts with methane to form a molecule in which two carbon atoms are bonded together, ethene, C_2H_4 (structure 6):



Ethene absorbs ultraviolet radiation and loses hydrogen to form ethyne ($\text{HC}\equiv\text{CH}$). Because ethene has a high probability of reacting with the solar radiation, it forms ethyne almost as soon as it is produced and so there is very little of it around. The observed abundance of ethene on Jupiter is lower than those of ethane ($\text{H}_3\text{C}-\text{CH}_3$) and ethyne.

Ethyne is quite abundant on Jupiter. This is because when it is broken up by radiation, the products rapidly react with hydrogen to re-form ethyne:

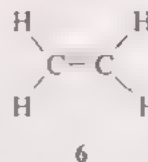


On Titan, however, the major constituent of the atmosphere is not hydrogen but nitrogen, and so the radicals $\text{HC}\equiv\text{C}$ do not reform ethyne but go on to produce larger molecules. One of the simplest ways they can do this is illustrated by Equations 7.12 and 7.13.



Successive reactions of this sort can produce very long molecules with hundreds of linked carbon atoms.

Write an equation for the formation of the next molecule in the series (that is one with eight carbon atoms).



From now on we shall meet carbon atoms that are bound to each other by more than one chemical bond. We shall show the number: $\text{C}-\text{C}$ is one, $\text{C}=\text{C}$ two, and $\text{C}\equiv\text{C}$ three.

It is not known how far the production of carbon chains proceeds on Titan, but it is possible that some of the longer-chain molecules may be responsible for the haze observed in the atmosphere, this haze thus being a photochemical smog. The formation of very long chains may be inhibited by the presence of hydrogen atoms in the atmosphere.

□ Why would hydrogen atoms affect reactions such as 7.12 and 7.13?

■ Hydrogen atoms are produced in these reactions. If hydrogen atoms are present from some other source these will promote the reverse reactions.

Interest in the formation of long-chain molecules is fuelled by the possibility that some of these may be precursors for the molecules needed to produce life. It is unlikely, however, that life ever developed on Titan.

Summary of Section 7.6 and SAQs

- 1 In the outermost layers of the atmospheres, absorption of ultraviolet radiation plays an important role in the chemistry. This can lead to the presence of molecules that would not be expected were the atmospheres in chemical equilibrium.
- 2 Absorption of ultraviolet radiation by ammonia can lead to a series of reactions in which nitrogen, N_2 , is a product. This may be the explanation for the abundance of N_2 in Titan's atmosphere, but it is also possible that nitrogen was present as N_2 rather than NH_3 when Titan formed.
- 3 Methane, CH_4 , can undergo a series of reactions that lead to the formation of molecules with two or more carbons linked together.
- 4 Longer carbon chains are found on Titan than on Jupiter because there is little hydrogen (H_2) to react with the smaller molecules.

SAQ 7.15 (Objective 7.10) Is there any evidence that methane, CH_4 , absorbs ultraviolet radiation and undergoes reactions in the atmosphere of Neptune similar to those described here?

SAQ 7.16 (Objective 7.4) Photochemical reactions of NH_3 and CH_4 (ammonia and methane) are important in the atmospheres of the giant planets, but reactions of H_2O (water) are not. Suggest why this might be so.

7.7 Overview of the giant planets

The giant planets are very different worlds from the terrestrial planets. Jupiter and Saturn are composed mostly of hydrogen and helium and no surface has been detected. Current models indicate a dense core of rocky materials, or rocky and icy materials, which occupies only a small fraction of the planet's volume. A thick layer of metallic hydrogen surrounds the core, and around this is fluid molecular hydrogen. Uranus and Neptune are believed to consist mainly of rocky and liquid icy materials, the latter being ionized to a considerable extent, surrounded by a deep atmosphere of mainly hydrogen and helium. The metallic hydrogen and liquid icy layers can account for the observed magnetic fields.

Jupiter, Saturn and Neptune are known to have hot interiors because they give out more energy than they receive from the Sun. In Jupiter, this is thought to be due to accretional heat and heat of initial differentiation. In Saturn, a further contribution is thought to come from the ongoing separation of liquid helium from metallic hydrogen. Uranus may have a hot interior, but this is not manifest in an excess of energy given out.

Although the atmospheres of the giant planets are mainly hydrogen and helium, several other molecules have been detected, the main ones being ammonia, NH_3 , and methane, CH_4 . What we see when we look at these planets is the topmost cloud layer. On Jupiter and Saturn this is composed of ammonia droplets, and on Uranus and Neptune it consists of droplets of methane. On Jupiter and Saturn other cloud layers are thought to exist deeper down, notably of solid ammonium hydrogen sulphide (NH_4HS) and water ice. These layers are predicted by calculations that assume chemical equilibrium and relative solar abundances of the elements; they have not been observed, however.

The main constituent of Titan's atmosphere is nitrogen, N_2 , and there is a haze of hydrocarbons.

The atmospheric temperatures in the measurable layers of all four planets decrease with distance from the centre initially, giving a troposphere, and then increase in the thermosphere.

Wind velocities, measured by tracking cloud features and comparing their rotation speed with that of the centre of the planet, show an alternating pattern from equator to pole. At the equator very high wind speeds are found, in the direction of rotation for Saturn and in the opposite direction for Neptune.

Hurricane-like storms also appear in the atmospheres of these planets and are longer lived than on Earth. The most famous is the Great Red Spot of Jupiter.

Above the clouds, molecules have been observed that would not be expected at chemical equilibrium. These are the result of photochemical reactions.

We now turn from the largest constituents of our planetary system (the giant planets) to examine, in the final chapter of this book, some of the smallest: meteorites and the grains within them, and comet dust.

Objectives for Chapter 7

After studying Chapter 7 (and any associated audio, video and TV material) you should be able to:

- 7.1 Give brief definitions of the terms, concepts and principles listed at the end of the Objectives.
- 7.2 Describe the (generally accepted) structures of the giant planets and of Titan.
- 7.3 Show how various observational data support (or otherwise) a given model for the structure.
- 7.4 Describe the main features of the chemical composition of the atmospheres of the four giants and of Titan.
- 7.5 Use information on temperature and spectroscopic selection rules to discuss how a particular molecule or atom might be detected.
- 7.6 Explain the principles by which the compositions of the atmospheres below the top cloud deck are predicted.
- 7.7 Describe the vertical variation of temperature in the atmospheres of the four giants and Titan.
- 7.8 Explain how wind velocities are measured, and describe the main features of the atmospheric circulation on the giant planets.
- 7.9 Describe the heat sources available on the four planets to drive the atmospheric circulation.
- 7.10 Explain how sunlight plays an important role in determining the composition of the atmospheres of Jupiter, Saturn, Uranus, Neptune and Titan, and give chemical equations for some of the reactions involved.

List of scientific terms, concepts and principles used in Chapter 7

Term	Page	Term	Page	Term	Page
belts	189	hydrocarbons	198	wind speed/velocity	187
chemical equilibrium	184	metallic bonding	174	zones	189
equilibrium constant	184	metallic hydrogen	173		
Great Red Spot	190	Neptune's Dark Spot	193		

Answers to Exercise on Neptune's Dark Spot

Question 1

In the article, there are three references to Saturn and two to Uranus. For Uranus we learn "Voyager 2 found Uranus' calm, blue outer layers nearly featureless" and "why just on Jupiter and Neptune – why not on Saturn and Uranus?". For Saturn the article is more ambiguous. We learn in paragraph 1 that "Only three minor, low-contrast, oval weather systems were found in Saturn's less-turbulent atmosphere" but towards the end of the article there is a paragraph stating "We must also include the storms of Saturn in this comparative study." Taking these statements together, we arrive at the conclusion that no comparable features are seen, but that this is not because similar storms do not arise on Saturn. There are large oval storm systems on Saturn but they are not strongly coloured. This article was written before the reappearance of the White Spot on Saturn in 1990 (Subsection 7.5.3).

Question 2

Paragraph 3 deals with the investigators' surprise. The reasons given are that very little sunlight is received on Neptune, and its lower mass relative to Jupiter. Large amounts of energy are involved in storms such as the Dark Spot (think of the damage wrought by hurricanes on Earth!). This energy must come either from the internal heat of the planet or from the Sun.

Question 3

The fifth paragraph mainly deals with the circulation near the Dark Spot. We learn that its boundaries are not as sharp as in the Great Red Spot and that what keeps both spots' shape is the surrounding winds. The surrounding clouds rise and flow over the storm without being consumed, and yet paragraph 2 suggests that the Great Red Spot is formed by larger eddies consuming smaller ones. Finally, there is no evidence of the Dark Spot being on the edge of a strong wind jet, although we have seen that this is where features occur on Jupiter.

Question 4

The Great Dark Spot strongly absorbs different wavelengths of light from the ones absorbed by the Great Red Spot. This suggests a different chemical composition for the coloured substances.

Chapter 8

Meteorites and comets

Prepared for the Course Team by Ian Wright and Colin Pillinger

Contents

8.1	Introduction	204
8.2	An overview of meteorites and micrometeorites	204
8.2.1	Cosmic debris	204
8.2.2	Meteorites, meteoroids and meteors	207
8.2.3	Meteorite origins	208
	Summary of Section 8.2 and SAQs	209
8.3	Allende and Murchison – cosmic sediments	210
8.3.1	Primitive materials from the birth of the Solar System	210
8.3.2	Refractory inclusions – hosts of the first minerals to condense in the solar nebula	211
8.3.3	Hydrated minerals and organic compounds – indicators of low-temperature processes	215
8.3.4	Silicon carbide – refractory material from beyond the Solar System	218
	Summary of Section 8.3 and SAQs	219
8.4	Juvinas and Cape York – highly processed materials	220
8.4.1	Melting processes in asteroids	220
8.4.2	Juvinas – a basaltic partial melt	221
8.4.3	Cape York – a sample from the core of an asteroid	223
	Summary of Section 8.4 and SAQs	225
8.5	EET A79001 – a meteorite from Mars	225
8.5.1	Collection in Antarctica	225
8.5.2	Why do we think that EET A79001 comes from Mars?	225
8.5.3	Nitrogen on Mars	228
8.5.4	Evolution of the Martian atmosphere – the role of water	230
	Summary of Section 8.5 and SAQs	232
8.6	Micrometeorites – samples of comets	233
8.6.1	The Giotto space mission to Halley	233
8.6.2	Interplanetary dust	233
8.6.3	Laboratory analyses of cometary dust	234
	Summary of Section 8.6 and SAQs	237
8.7	Conclusions	237
	Objectives for Chapter 8	237

8.1 Introduction

We turn now from a study of the largest planetary bodies in the Solar System to some of its smallest constituents. **Meteorites** are pieces of extraterrestrial material that fall to the Earth from the sky. In the past, these objects were the subject of much curiosity – some meteorites have been revered and worshipped, while others have been collected and used for making tools. Today, meteorites are studied intensively by scientists using every manner of analytical instrumentation. Video sequence 8 presents one example, and you are asked to watch this during your study of Subsection 8.3.2.

The purpose of this chapter is not to describe every aspect of the study of meteorites, but to select a few examples of particularly interesting samples. We will try to convey how meteorite analyses contribute to our understanding of the Solar System. In Section 8.2 there is a brief introduction to meteorite fall statistics, sample collection, and classification. Section 8.3 details what can be learned from the study of the most primitive meteorites, with particular reference to processes that took place within the solar nebula during the period when it was in the process of forming the Solar System. Section 8.4 describes meteorites that were formed as a result of heating and melting events early in the history of the Solar System, while in Section 8.5 a meteorite with a presumed Martian origin is discussed. Micrometeorites, and their relationship to comets, are described in Section 8.6.

8.2 An overview of meteorites and micrometeorites

8.2.1 Cosmic debris

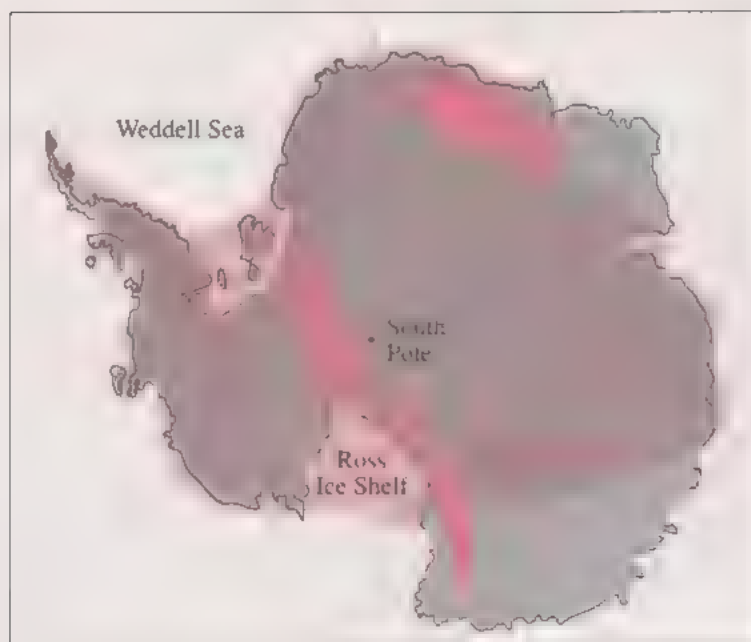
The collection of a meteorite sometimes takes place after it has been observed to travel through the Earth's atmosphere. This is known as a meteorite **fall**. Events like this are rarely witnessed – only about 6–8 falls are recorded each year. In contrast, several hundred meteorite fragments are collected each year without being seen to fall; these are the meteorite **finds**. How many meteorites arrive at the Earth annually? It is known from continuous observations of the night sky that 20 000 bodies with masses in excess of 0.1 kg should reach the Earth's surface each year. Even of those that fall on land, it is clear that the majority of large meteorites are never recovered.

Estimates of the total number of meteorites present in various collections around the world vary from 4 000–10 000.

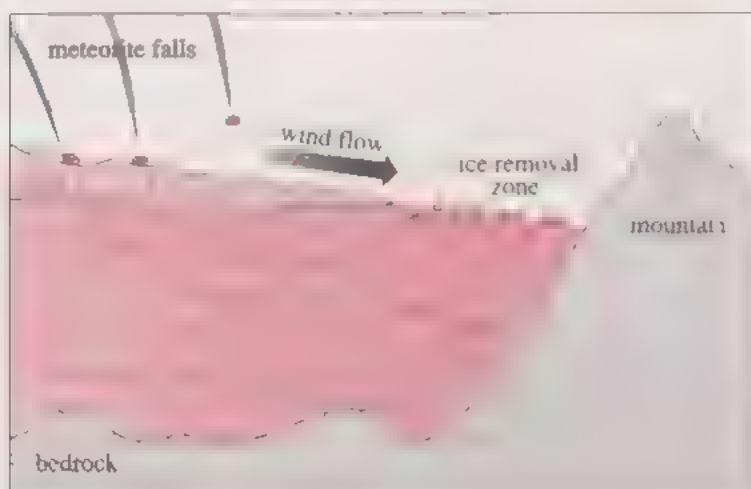
Why do you think there is so much uncertainty with this number?

■ Sometimes a large number of separate fragments may arise from an *individual* meteorite fall. This is known as a meteorite shower.

Meteorites can be found in any continental location (Figure 4.8 showed an example), but various factors affect the chances of recovery. For instance, areas of dense vegetation are not good hunting grounds for meteorites. In contrast, meteorites that fall in relatively barren environments, such as hot deserts, have a greater potential for collection. Antarctica is an even better place for finding meteorites, since glacial processes sometimes act to concentrate the samples (see Figure 8.1, and the TV programme *Cosmic recycling*). Dedicated collection parties frequently visit places like Antarctica in order to collect materials for scientific research. Plate 2.46 shows some examples of meteorite collection.



(a)



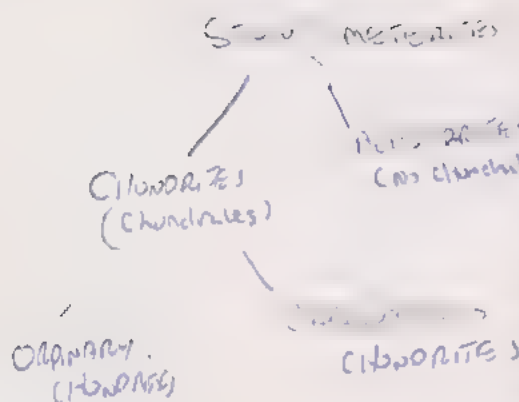
(b)

Figure 8.1 (a) Map of Antarctica showing locations of meteorite finds (red areas). (b) Schematic cross-section through the Antarctic ice sheet showing how ice movement acts to concentrate meteorites from a wide area in which they fall. Where a glacier abuts a mountain range the ice flow (red arrows) is diverted upwards. Strong winds blowing from the South Pole cause the ice to be removed, leaving behind the meteorite samples.

Meteorites can be broadly classified as irons, stones or stony-irons. As their names might suggest, irons (or **iron meteorites**) are predominantly composed of iron metal while stones (or **stony meteorites**) are comprised mostly of silicate minerals. **Stony-iron meteorites** are mixtures of metals and silicates. Plate 2.47 shows a variety of different meteorite types.

In order to learn something of the origins of meteorites it is important to know the relative proportions of irons, stones and stony-irons that arrive at the Earth. For instance, which meteorites are the most common and thus most typically represent the source? Studying the meteorite finds cannot answer this question. This is because a piece of metal weighing several kilograms is obviously unusual and more likely to be collected as a find than a silicate meteorite, which may look like any other piece of rock. In order to properly assess which sorts of meteorites are most common it is necessary to consider only those samples *observed to fall*. Using this criterion, irons constitute less than 5% of all meteorites and stony-irons amount to about 1%. Note that if we had considered only the meteorite finds, irons would appear to constitute 40% of all meteorites.

Stony meteorites are the samples that fall to Earth most frequently. Of these, more than 90% are **chondrites**, i.e. samples that generally contain **chondrules**.



☐ Can you remember what chondrules are?

☒ Chondrules are globules of silicate minerals, up to a few millimetres in size (see Subsection 2.3.2). They have been interpreted as rounded particles of rapidly cooled silicate melt formed by condensation, or re-melting of dust, in the solar nebula. (Alternatively some chondrules may be formed at the surfaces of planetesimals as the result of impact melting.)

Most chondrites belong to a group known as the **ordinary chondrites**. The term 'ordinary' is used because they are the most common sort of meteorite. Ordinary chondrites contain varying proportions (5–15%) of iron–nickel metal, a factor which, along with the presence of chondrules, makes them distinctive from terrestrial rocks. Another important chondrite group is represented by the **carbonaceous chondrites** (Subsections 2.2.3 and 2.3.2), these are carbon-rich meteorites that contain organic compounds such as amino acids. The so called **achondrites** are stony meteorites that do not contain any chondrules, these have **generally been heated to their melting temperatures, and formed by crystallization of magmas on their respective parent bodies**. Some achondrites are mechanical mixtures of melted fragments and pieces of chondrites. These meteorites formed at the surface of a parent body where, early in the history of the Solar System, impact processes assembled a variety of different materials together.

A typical stony meteorite has a mass of 1–10 kg (diameter 0.1–0.2 m), while the mass of an iron is generally in the range 10–40 kg. Although meteorites come in all shapes and sizes, there are many more small samples than large ones. Meteorites with masses greater than 100 kg are rather rare (the largest single meteorite weighs 5.5×10^4 kg). The smallest, dust-sized materials (millimetre-sized, or less) are known as **micrometeorites** and have masses of less than 1×10^{-6} kg. Samples of this nature are constantly falling to Earth. Indeed, a person who spends a few hours outside during the day is likely to be hit by at least one extraterrestrial particle (albeit only 0.01 mm in size!). This constant rain of material to Earth amounts to about 1×10^7 kg yr⁻¹. This sounds a lot, but the **total annual flux of extraterrestrial material, which includes bodies larger than micrometeorites, is estimated to be about 1×10^8 kg yr⁻¹**. It should be noted that **since the mass of the Earth is 5.98×10^{24} kg – i.e. more than 16 orders of magnitude larger – the yearly input of extraterrestrial material does not cause any obvious changes to the physics or chemistry of our planet as a whole.**

ITQ 8.1 Assuming that the flux of extraterrestrial material has remained constant since the Solar System formed, calculate the mass and volume that this extraneous matter has added to our planet over the last 4 billion years. If all the extraterrestrial materials were available now as a surface layer across the whole of the planet, how deep would this be? (Assume a density for the incoming material of 1.5×10^3 kg m⁻³.)

What would it feel like to be hit by a micrometeorite? Let's consider how fast these objects are travelling. In space they travel at speeds typically of 30–42 km s⁻¹. Fortunately the Earth's atmosphere acts to retard projectiles with masses of less than a few kilograms (equivalent to a body of 0.1 m diameter). These smaller bodies fall to Earth under the influence of gravity opposed by atmospheric drag and have impact speeds of less than 5×10^{-3} km s⁻¹. In fact, the speed of fall of a micrometeorite is nearer 1×10^{-6} km s⁻¹. Much larger objects, on the other hand, arrive at the Earth's surface with a significant fraction of the speed they had in space.

☐ What physical effects will large meteorites exert on our planet?

■ You saw from Chapter 4 that, following the collision of cosmic debris with various planetary bodies, impact craters are produced which may be many kilometres in diameter.

A logical question to ask at this point is whether such 'cosmic wanderers' could, in the future, be detected and thereby prevented from impacting on the Earth? A plausible mechanism of interception remains in the realm of science fiction although a number of research groups have addressed the issue. Detection, on the other hand, is somewhat more certain. We know that the large, potentially destructive bodies originate in the asteroid belt (see Subsection 2.5.4). Recall that the main belt of asteroids is a collection of planetesimals, debris and dust between the orbits of Mars and Jupiter (Section 1.3). The orbits of some asteroids are believed to evolve to the point where they intersect that of the Earth (a consequence of gravitational perturbations by Jupiter). Collectively these are known as the Earth-crossing asteroids (Section 1.3). It is these asteroids which have the potential to impact on our planet. About 100 Earth crossing asteroids are currently known, but detection has been limited to those of relatively large size (>0.5 km diameter). The orbital parameters of smaller bodies are generally unknown. Not surprisingly, unexpected close encounters happen from time to time. For instance, on 23 March 1989, a previously unknown asteroid of 0.3 km diameter passed within 6×10^5 km of the Earth. Indeed, it was detected only as it moved away from us. Although estimates of the total number of Earth-crossing asteroids have been put at 1 500, recent work has shown that as many as 50 small asteroids may pass between the Moon and Earth each day!

8.2.2 Meteorites, meteoroids and meteors

The term meteorite is used to denote a sample of material that has fallen to Earth and has subsequently been collected as a coherent mass. While still in space the same body is referred to as a **meteoroid**. Meteoroids are formed from asteroids as a result of collisional processes. We know that they come from asteroids since, in three well documented cases, meteoroids which became meteorites were photographed during their fall, using networks of cameras. From the photographic evidence it has been possible to assess the meteoroids' orbits, which can be traced back to the asteroid belt (see Figure 8.2).

For those meteoroids that become meteorites the brief atmospheric passage is quite eventful. People who have witnessed falls can attest to the spectacular visual display and accompanying detonations. At an altitude of about 120 km frictional heating by the atmosphere begins to melt the surface of an incoming meteorite. For a body of >10 mm in diameter, only the outer layers at any instant are heated sufficiently to cause melting and vaporization. This heating of the surface results in the formation of a glassy **fusion crust**, which is a further diagnostic indicator of a meteorite. During fall, molten material is lost from the surface by a process called **ablation**. Perhaps 50% of the total mass of a meteoroid is lost this way.

During their fall to Earth the very smallest meteoroids, i.e. **micrometeoroids**, become (not surprisingly) **micrometeorites**. Unlike larger samples, micrometeoroids, which are of roughly millimetre-sized dimensions (i.e. in the range 10–0.01 mm), become *totally* molten during their fall. Some of these vaporize completely while others survive to form **cosmic spherules**, which can be collected at the Earth's surface. Ordinarily these spherules go unnoticed. However, they are found in abundance in oceanic sediments (they were first identified in Pacific Ocean clays during the Challenger expeditions of the 1880s). Plate 2.49 shows a picture of cosmic particles extracted from a deep-sea sediment. Curiously, the very smallest micrometeoroids (<0.01 mm in size) do not melt during their fall to Earth. This is because they are decelerated at sufficiently high enough levels, where the atmosphere is more rarified, and so avoid melting by frictional heating.

Effectively, meteoroid is just a word for a small asteroid. There is no generally-accepted size limit but 100 m diameter is about right.

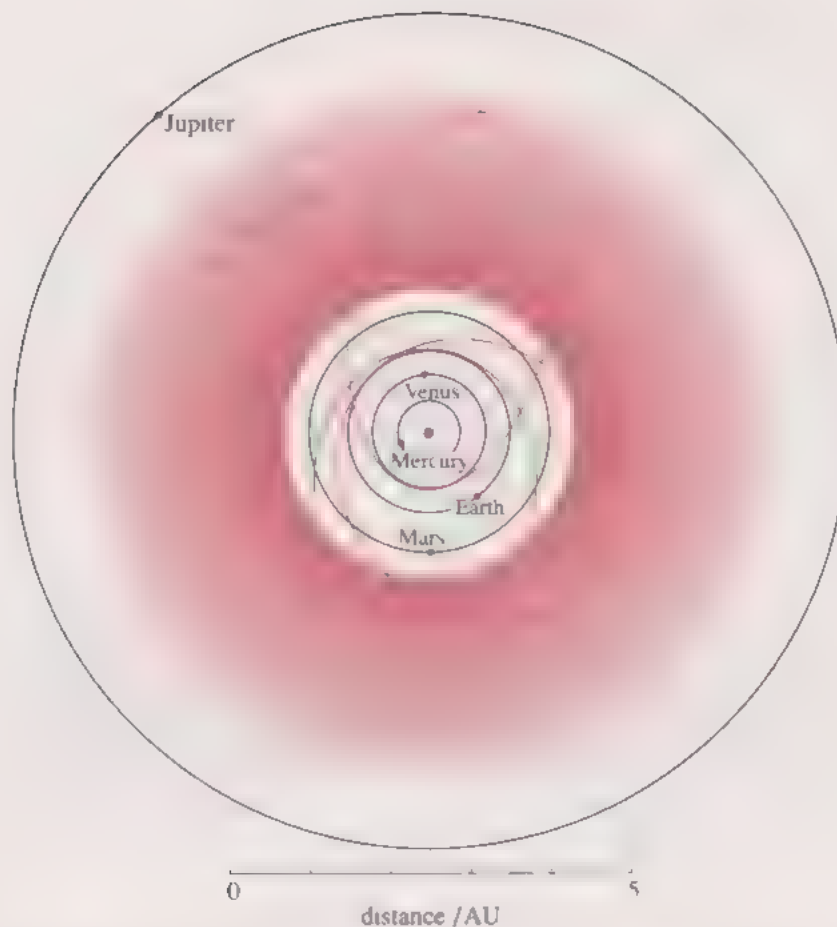


Figure 8.2 The orbits of three meteorites (shown in red) have been determined from data recorded by photographic networks (one in Czechoslovakia in 1959, one in the USA in 1970 and one in Canada in 1977). In these three cases the orbits can be traced back to the asteroid belt.

In addition to the heating episode, the atmospheric gases surrounding an incoming meteoroid become ionized and form a plasma. As the meteoroid travels through the atmosphere a streak of light is thus produced, known as a meteor, or, if bright enough, a **fireball**. Typically, fireballs become extinguished at an altitude of about 10–30 km. For large meteoroids the fireball can be seen in daylight, as can the accompanying trail of ablated material (which has the appearance of smoke). For smaller bodies, a meteor trail is only visible at night, when it may be referred to as a shooting star. Practically all meteors are formed from incoming micrometeoroids.

As a meteoroid descends to Earth, atmospheric drag causes the body to slow down. Eventually, surface melting ceases. For a meteoroid of >0.1 m diameter the interior remains relatively cool during atmospheric entry; in fact, immediately following their fall, meteorites that have been broken open have sometimes been observed to form frost on their inside surfaces.

8.2.3 Meteorite origins

You have seen how meteorites can be distinguished, on the basis of their principal constituents, as irons, stones etc., and you have learned that meteorites are thought to originate in the asteroid belt. In Subsection 1.3.1 you learned about different sorts of asteroids. If we could relate the various meteorite types to asteroid groups we would effectively establish their origin, and be able to study the asteroids in great detail using conventional laboratory-based analytical techniques.

□ Can you remember how asteroids are classified?

■ Asteroids can be classified as C-type and S-type, using a technique known as photometry (which you met in Book 1, Subsection 2.3.2). Essentially this entails measuring the brightness of reflected sunlight at different wavelengths (see Subsection 1.3.1).

Meteorite surfaces can also be analysed using a form of photometry. In this way we find that the C-type asteroids most closely resemble carbonaceous chondrites, while the S-type asteroids represent stony-irons. A further group, the M-type asteroids, look like iron (metallic) meteorites. The agreement between results obtained from the photometry of asteroids and analyses of meteorites is circumstantial evidence for a relationship between the two. Unfortunately we have not yet found a match between the most common group of meteorites, the ordinary chondrites, and any asteroid group. This somewhat embarrassing problem continues to drive asteroid/meteorite research!

Micrometeoroids do not have a single origin. Some are formed by collisional processes in the asteroid belt, just like their larger counterparts. Others are ablation products formed in the Earth's atmosphere from the passage of meteoroids. However, the majority of micrometeoroids represent fresh debris liberated by comets near perihelion (Table 1.1d, footnote a). If large quantities of dust are released from a comet this may end up going into a specific orbit, in which case the dust will be encountered by the Earth at the same time each year. When this happens the incidence of meteors increases – a **meteor stream** is witnessed. A table in *Images of the Cosmos* lists meteor streams and the source comets associated with them.

In addition to asteroids and comets, a few meteoroids have origins on planetary bodies. Several meteorites are known to have come from the Moon. These were ejected when the Moon was itself bombarded by a projectile (meteoroid, or comet). How can we be so sure about their lunar origin? One way is to compare the chemistry of these samples directly with materials returned from the Moon (by US and Soviet spacecraft during 1969–76). A further group of meteorites (currently represented by 10 specimens) has been recognized as unusual for over a century – these samples are now thought to have originated on Mars, and you will learn more about them in Section 8.5.

The value of meteorite research lies in the fact that we have pieces of extraterrestrial material available for study in the laboratory. As we have already seen, meteorites come from asteroids, the Moon and Mars. Some of the micrometeorites come from comets. These materials arrive on Earth at no financial cost. They do not always represent the ideal sample, having been heated during atmospheric passage, fragmented during impact, and weathered and/or contaminated before collection. However, like good tourists, these samples acquire a record of their travels. This often enables an effective reconstruction of their histories.

Summary of Section 8.2 and SAQs

- 1 Meteorites can be broadly classified into stones, irons and stony-irons. The stones are by far the most common variety.
- 2 The fall of a meteorite is rarely witnessed. In contrast, many samples are collected each year as finds. In some cases these are retrieved by dedicated collection parties working in deserts (such as Antarctica).
- 3 There is a wide range in the size of extraterrestrial material which impacts on the Earth. Micrometeorites are fine dust of millimetre size, or less. Asteroids of kilometre dimensions have left their mark on the Earth's surface as spectacular impact craters. A typical stony meteorite has a mass of 1–10 kg (0.1–0.2 m diameter).
- 4 Meteorites have a number of sources – asteroids, the Moon and Mars. Micrometeorites arise as a result of collisional processes in the asteroid belt, and from comets near perihelion.

SAQ 8.1 (Objective 8.2) A meteorite has been found which is composed of a mixture of iron and silicates. Within the silicate portion there are fragments of materials recognizable as chondrites and achondrites. The meteorite has obviously been subjected to intense shock. Consider how it might have formed.

SAQ 8.2 (Objective 8.2) What is the evidence that most meteorites come from the asteroid belt?

SAQ 8.3 (Objective 8.3) 20 000 meteoroids of >0.1 kg arrive at the Earth each year. On average, what is the frequency with which materials of this size would be encountered in a small village (an area of, say, 50 km^2)?

8.3 Allende and Murchison – cosmic sediments

8.3.1 Primitive materials from the birth of the Solar System

Meteorites are generally given the name of their collection site, regardless of whether they represent a fall or find. Thus, the meteorite which fell in 1969 near to the village of Pueblito de Allende, in Mexico, is known as Allende. Like many other meteorites, Allende fragmented during descent through the atmosphere, resulting in a shower of stones. More than 2 000 kg of material was subsequently gathered over an area of about 300 km^2 . Murchison, a shower of 500 kg, also fell in 1969, in Victoria, Australia. Meteorites like Allende and Murchison have been studied intensively. You will see in this section how some of the components of these meteorites appear to document changes in formation conditions within the solar nebula. You will also learn that some of their mineral grains and organic compounds pre-date the formation of the Solar System; by studying these materials it is possible to gain an insight into various galactic processes, such as stellar nucleosynthesis, condensation around stars and interstellar chemistry.

Allende and Murchison are both carbonaceous chondrites. As you saw in Subsection 2.2.3, meteorites of this nature have chemical compositions that, in the case of all but the most volatile elements, match that of the Sun (recall that an element is regarded as volatile if it is itself volatile, or if it forms volatile compounds). This similarity with the Sun demonstrates that carbonaceous chondrites are very **primitive** materials, which have not had their chemistry changed by secondary processes. Note that certain other groups of meteorites, and indeed terrestrial crustal rocks, do not show such a good correspondence in **chemical composition with the Sun**.

When we look in detail at carbonaceous chondrites we find that there are various groups that can be distinguished on the basis of slightly different chemical compositions. Allende is a type 3, or C3, carbonaceous chondrite, whereas Murchison is a C2. The meteorites which most closely resemble the Sun are the C1 carbonaceous chondrites (Plate 2.47c). C1 samples are chemically the most primitive of all; a progression away from primitiveness is observed in samples of successively higher types. This can be understood by looking at Figure 8.3; here are plotted the abundances of certain elements found in the silicate and oxide minerals of Murchison and Allende. The elements are arranged from left to right according to volatility. It would be meaningless to plot absolute concentrations in Figure 8.3 since carbonaceous chondrites contain variable amounts of water, as hydrated minerals and organic compounds (which you will learn more about in Subsection 8.2.3). Instead, the elemental data are normalized to a non-volatile element – in this case silicon (so, for instance, the Ca/Si mass ratio is used). These are then plotted with respect to the equivalent ratios found in C1 meteorites. This may all seem a little complicated until you realize that in Figure 8.3 the chemical composition of a typical C1 meteorite would plot as a horizontal line, at a value of 1.0.

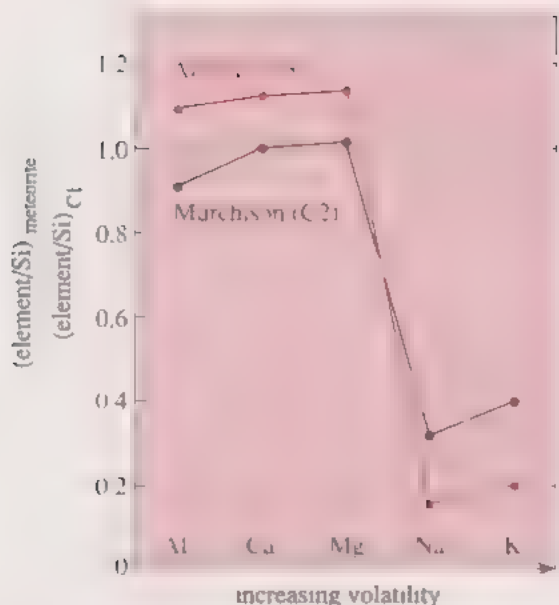


Figure 8.3 Plot of elemental concentrations of aluminium, calcium, magnesium, sodium and potassium (all normalized to silicon) in Allende (C3) and Murchison (C2), relative to data from C1 meteorites. The most primitive samples are represented by the C1 data, i.e. a horizontal line at a value of 1.0.

Refer back to Table 2.1 and assess the volatilities of aluminium and calcium relative to other elements.

■ Corundum (Al_2O_3) and perovskite (CaTiO_3) are some of the first minerals predicted to condense (at temperatures in excess of 1 600 K) from a cooling gas of solar composition. Al and Ca are thus refractory elements. For comparison, the volatile elements Na and K are incorporated into alkali feldspars and appear in the condensation sequence at < 1 000 K.

You can see from Figure 8.3 that the elements Al, Ca and Mg are somewhat more abundant (relative to silicon) in Allende (C3) than in the other two types. These elements also lie on roughly horizontal lines (other refractory elements also display this trend, although for simplicity these have been left off the figure), showing that refractory elements occur in approximately the same relative proportions to each other in C1–C3 carbonaceous chondrites (the Ca–Al ratio is virtually constant, for instance). The volatiles in Murchison and Allende (Na and K) are depleted with respect to the refractory elements. Note that the depletion is greater in the C3 sample than the C2. Thus, in going from C1 to C3 meteorites, the chemical compositions become progressively less like that of the Sun, i.e. less primitive.

Why have we not selected a C1 meteorite to study in detail? Unfortunately, meteorites which are the most primitive in terms of chemical composition have also suffered the complication of **hydrothermal alteration**. C1 meteorites were endowed with a large complement of volatiles, they contain about 20% by mass of H_2O , for instance. Transient heating events on the parent bodies of these meteorites (during impacts, for instance) have acted to mobilize volatiles such as water, which in turn have reacted with the more refractory components. The effect of this alteration is so intense that recognizing primary features is difficult – the bulk of the meteorites have been converted into complex hydrated minerals somewhat similar to those found in terrestrial clays.

ITQ 8.2 The chemical composition of C1 meteorites is considered to be the same now as it was when they were formed, even though the original minerals had been altered. What can you say about the nature of the alteration process which has affected C1 meteorites?

A detailed look at carbonaceous chondrites shows that the effects of hydrothermal alteration are most intense in C1 samples, less so in C2s and are virtually absent in C3s. Thus, although a C3 meteorite like Allende is not the most chemically primitive of samples, it is representative of the least-altered of carbonaceous chondrites. It is therefore a good place to commence with the study of meteorites.

Allende is composed of rounded chondrules and large irregular-shaped white inclusions set in a fine-grained, dark matrix. Some examples of chondrules are shown in Plates 2.47c and 2.48a and b. This collection of materials was assembled on a parent body as dust settled out from the solar nebula. This action can be likened to the deposition of a sediment – for this reason we think of carbonaceous chondrites as **cosmic sediments**. Be sure that you understand that there were no liquids present during the meteorite formation process. The water which was imparted to C1 and C2 meteorites settled out with the refractory dust as solid ice grains.

8.3.2 Refractory inclusions – hosts of the first minerals to condense in the solar nebula

The white inclusions in Allende contain refractory elements. Because of their chemical compositions they are frequently referred to as **calcium–aluminium-rich inclusions**, or **CAIs** for short. The mineralogy of these refractory inclusions matches very well with what is predicted to form at high temperatures in the condensation sequence (see Subsection 2.3.1). It would appear, therefore, that the inclusions contain some of the first minerals to form in the cooling solar nebula. Equally well, however, they may represent evaporation residues (i.e. aggregations of minerals that lost all traces of volatile materials through heating and evaporation). An alternative view holds that the CAIs could in some way represent pre-solar materials that were never heated significantly in the solar nebula. In this case, formation of these materials is considered to have taken place before the solar nebula formed, in environments such as the outflow of supernovae.

Is there a way that we can try and distinguish between the competing theories for the origins of refractory inclusions? One way is to measure the **stable isotope compositions** of the constituents. Such compositions are discussed in Box 8.1.

Box 8.1 Stable isotope compositions

You saw in Subsection 3.5.2 that certain elements have radioactive isotopes. Potassium, for instance, has a radioactive isotope, $^{40}_{19}\text{K}$, which decays to give $^{40}_{18}\text{Ar}$. But potassium also has two further isotopes, $^{39}_{19}\text{K}$ and $^{41}_{19}\text{K}$, which are not radioactive. These are called **stable isotopes** because they do not decay. Most elements are comprised of mixtures of stable isotopes; oxygen has three $^{16}_8\text{O}$, $^{17}_8\text{O}$ and $^{18}_8\text{O}$. All the isotopes of oxygen contain 8 protons – this is the atomic number. However, the isotopes differ in their number of neutrons; $^{16}_8\text{O}$, $^{17}_8\text{O}$ and $^{18}_8\text{O}$ have 8, 9 and 10 respectively.

In astronomical entities, such as stars, the variations in oxygen isotope compositions can be quite large and so the differences are described in terms of their $^{16}\text{O}/^{17}\text{O}$ and $^{16}\text{O}/^{18}\text{O}$ ratios. For meteorite samples (or indeed, terrestrial rocks) this is not practical because the variations in oxygen isotope composition are relatively minor – tremendous significance may be placed on deviations of less than 1%. In order to describe these small variations it is convenient to express the *differences* in isotope compositions, rather than the absolute ratios. This is done by relating the sample ratio to a reference point. In the case of oxygen, the water which makes up the Earth's oceans is usually used as a reference since this is essentially homogeneous with respect to isotope composition, having $^{16}\text{O}/^{18}\text{O}$ of 499 and $^{16}\text{O}/^{17}\text{O}$ of 2 690. On the other hand, the oxygen that makes up the rocks of the Earth's crust has a somewhat variable isotope composition (dependent upon many factors, including formation conditions, temperature, etc.). A typical basaltic rock may have $^{16}\text{O}/^{18}\text{O}$ of 496 and $^{16}\text{O}/^{17}\text{O}$ of 2 682. The difference between

Recall that the chemical symbol (O here) implies the atomic number subscript, which is thus often omitted.

the oxygen isotope composition of ocean water and rocks is relatively small but easily and reproducibly measurable using an instrument known as a mass spectrometer (video sequence 8).

It is a convention to relate the stable isotope composition of an unknown sample to that of a reference material. Thus if R_s is the isotope ratio of the sample and R_r is the corresponding ratio in the reference, the difference is

$$\frac{R_s - R_r}{R_r} = \frac{R_s}{R_r} - 1$$

In this formula, we use the ratio of the rare to the abundant isotopes since, in effect, we are interested in variations of the rare isotopes. Because these differences are very small, they are usually quoted in parts per thousand (known as parts per mil), abbreviated to ‰. (Comparing this with parts per hundred, i.e. per cent (%), you can see that 1% = 10‰.) The calculated result, δ (delta, in ‰), is called the differential isotope composition, or more commonly the δ value.

$$\text{Thus, } \delta \text{ (in ‰)} = \left(\frac{R_s}{R_r} - 1 \right) \times 1000 \quad (8.1)$$

Since oxygen has three stable isotopes there are two different delta values, $\delta^{18}\text{O}$ and $\delta^{17}\text{O}$, defined relative to the abundant isotope, ^{16}O . Thus, for a $\delta^{18}\text{O}$ value, R_s and R_r refer to $^{18}\text{O}/^{16}\text{O}$ ratios. Similarly, for a $\delta^{17}\text{O}$ value, R_s and R_r refer to $^{17}\text{O}/^{16}\text{O}$ ratios. For the basaltic rock referred to above, inserting values into Equation 8.1 we get:

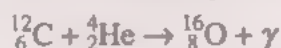
$$\delta^{18}\text{O} = \left(\frac{1/496}{1/499} - 1 \right) \times 1000 \approx +6‰$$

Since δ values can be either positive or negative numbers we explicitly give each of them a sign – thus, in the above example, $\delta^{18}\text{O} = +6‰$, rather than just 6‰. It can be shown that for this basaltic rock, $\delta^{17}\text{O}$ is $\approx +3‰$, equivalent to about half of the value of $\delta^{18}\text{O}$. The factor of about 0.5 difference between $\delta^{17}\text{O}$ and $\delta^{18}\text{O}$ holds true for almost all oxygen-containing materials on Earth. The reason for the relationship, which can be related to the difference in mass between the isotopes, is a consequence of the process known as isotope fractionation. This occurs when isotopes become separated, or fractionated, from each other (as can happen during chemical reactions or during physical processes such as diffusion).

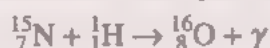
11Q83 What is the $\delta^{17}\text{O}$ and $\delta^{18}\text{O}$ of ocean water? What is the $\delta^{17}\text{O}$ and $\delta^{18}\text{O}$ of pure ^{16}O ?

You should now watch video sequence 8

Returning to the CAIs, if the inclusions represent samples of pre-solar dust, we may expect them to have a variety of different oxygen isotope compositions. This is because the oxygen isotopes are synthesized in various ways. The isotope ^{16}O can be formed directly (Book 1, Subsection 3.4.2):



or within the CNO cycle (Book 1, Subsection 3.3.3):



The other isotopes of oxygen, ^{17}O and ^{18}O , are not formed by helium burning, but can be produced in the CNO cycle. Thus, the oxygen isotope composition of any dust that condenses around a star will be dependent upon the relative

Remember that O is a catalyst in the CNO cycle, and that the *combined* abundance of C, N and O remains unchanged.

influence of the ${}^{12}\text{C}$ reaction versus that of the CNO cycle which operates in that particular star. In other words, variable proportions of ${}^{16}\text{O}$, ${}^{17}\text{O}$ and ${}^{18}\text{O}$ are produced within different stars. Other astronomical events, such as novae and supernovae, produce yet further variations in oxygen isotope compositions. Since the pre-solar dust that formed the solar nebula was produced from any number of stars, or other explosive events, we would expect the oxygen isotope compositions of this material to be highly variable.

In contrast, if CAIs were formed during condensation in the solar nebula we might expect a fairly straightforward relationship of the oxygen isotope compositions of the different inclusions. This is because a hot, turbulent nebula environment will act to homogenize the isotope compositions of pre-solar gas and dust.

A convenient way to consider oxygen isotope data is on a plot of $\delta^{17}\text{O}$ versus $\delta^{18}\text{O}$ (Figure 8.4). On a diagram like this, the oxygen isotope composition of samples from the Earth lie on the line called the **terrestrial fractionation line (TFL)** which has a slope of about 0.5 and passes through the point 0,0 where ocean water would plot.

□ Why does the TFL have a slope of about 0.5?

■ Because of the relationship between $\delta^{17}\text{O}$ and $\delta^{18}\text{O}$ in isotope fractionation (see Box 8.1).

No matter what oxygen isotope composition an individual body (planetary or otherwise) has initially, *secondary* processes that result in isotope fractionation produce data that fall on lines of slope 0.5. When this phenomenon is observed it is a good indicator that materials are related in some way. Two lines of slope 0.5 that are *not* coincident, imply different *initial* oxygen isotope compositions.

A plot of $\delta^{17}\text{O}$ versus $\delta^{18}\text{O}$ for data from the Allende CAIs is shown in Figure 8.4. In contrast to samples from the Earth, the oxygen isotope composition of different minerals from CAIs define a line labelled D–G which has a slope of about 1. That the data fall on a line, rather than at random across the plot, tends to suggest that the minerals were all formed in a common event. As it happens, refractory minerals, like spinel (Table 2.1), fall close to the lower end of the line (D), while less refractory minerals plot close to G.

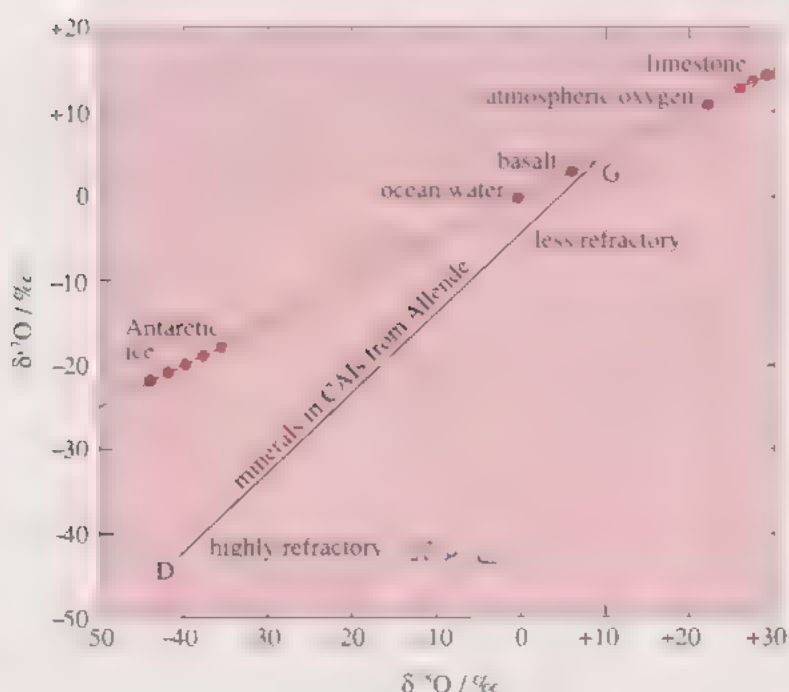


Figure 8.4 Plot of oxygen isotope compositions ($\delta^{17}\text{O}$ versus $\delta^{18}\text{O}$) of different minerals from refractory inclusions in Allende. The data fall on a line (D–G) with a slope of about 1. For reference, a line of slope about 0.5 and labelled TFL is also shown. Almost any sample from the Earth will have an oxygen isotope composition that falls somewhere on this line (a few examples are shown).

□ Remember that refractory minerals condense early in the condensation sequence. How can we explain the oxygen isotope data for the different minerals in CAIs?

■ The oxygen isotope composition of the reservoir from which the minerals condensed must have differed for spinel than for the less refractory minerals

The data have been explained in terms of a dust reservoir (D) and gas (G). The solar nebula started as a mixture of gas and very fine dust. It is proposed that of the minerals that we find in the CAIs, the first to form were somehow derived from the dust and reflect its initial oxygen isotope composition. This involved reprocessing of dust material, and so it is appropriate to refer to a dust *reservoir* rather than dust particles. As mineral formation proceeded, the isotope composition of the dust reservoir became modified by a process known as **isotope exchange** with the gas reservoir. The gas reservoir was much the larger, and so this exchange led to the dust reservoir becoming more like the gas reservoir. In consequence, the less refractory minerals that formed later, and now constitute part of the CAIs, lie closer to point G.

ITQ 8.4 Consider for a moment that the dust and gas reservoirs in the solar nebula *both* had oxygen isotope compositions at G in Figure 8.4. Where would the isotope compositions plot if we mixed small quantities of pure ^{16}O into the system? (Remember the δ values of pure ^{16}O from ITQ 8.3 and consider where this would fall on a plot like Figure 8.4.)

There are two important conclusions from the study of Allende CAIs. Firstly, not all of the oxygen in the solar nebula had the same initial isotope composition. This means that the solar nebula was not completely homogenized before condensation began (if this had happened the dust and gas reservoirs would be related to each other along a line of slope 0.5 in Figure 8.4, in the same way that different components from the Earth are related). Thus, meteorites contain some relicts of the pre-solar gas and dust which coalesced to form the solar nebula. Secondly, the oxygen isotope compositions of minerals in CAIs show a relatively simple pattern, i.e. evidence for only two major components with isotope exchange between them. This simple pattern was almost certainly established within the solar nebula, rather than before the solar nebula formed.

If we were to plot oxygen isotope data for the chondrules in Allende in Figure 8.4, we would find that they fall on a different line to that shown by CAIs. Although one end of the line would be coincident with D, the other would be slightly displaced from G. It would seem that in between the formation of CAIs and chondrules the oxygen isotope composition of the solar nebula had changed.

ITQ 8.5 What are the consequences for planetary bodies in general for a variation with time in the oxygen isotope composition of the solar nebula?

8.3.3 Hydrated minerals and organic compounds – indicators of low-temperature processes

In addition to chondrules and CAIs, Allende and Murchison are also comprised of a dark, fine-grained **matrix** (effectively, everything else!). The matrix is dark in colour, partly because it is fine-grained, but also because it contains carbonaceous materials. In Allende the matrix is predominantly an assemblage of mineral components such as olivine and sulphides (e.g. troilite, FeS). These minerals formed as a logical progression in the condensation sequence (see Table 2.1 on p. 35) – the fine-grained nature of the constituent grains may be due to fragmentation processes occurring in the solar nebula. In contrast to Allende, the

matrix of Murchison contains a large proportion of hydrated minerals. You can see from the condensation sequence that in a cooling gas of solar composition, hydrated minerals condense directly in the nebula at temperatures of 550–330 K.

What can you deduce from the presence of hydrated minerals in the matrix of Murchison (and their absence in Allende)?

■ The temperature of the nebula was in the region of 550–330 K when the matrix of Murchison was formed. That of Allende, on the other hand, must have formed at temperatures in excess of 550 K.

Not all of the features displayed by the matrix of Murchison can be explained by processes which took place during the condensation sequence. For instance, water-soluble substances, such as carbonates and sulphates, are found deposited as veins running through the fabric of the meteorite. It seems that Murchison, like C1 meteorites, has been affected by the action of fluids some time after it formed. Unlike the C1 meteorites, however, the alteration in Murchison is not that extensive – the primary constituents such as chondrules and CAIs are still recognizable. On the basis of geochemical evidence it can be deduced that the hydrothermal alteration took place at about 300 K.

Turning now to the carbonaceous materials, the matrix of Allende contains about 0.25% by mass of carbon, mainly in an amorphous form. Murchison, on the other hand, contains about an order of magnitude more carbon, not in amorphous form but as organic compounds (i.e. materials made up from the elements carbon, hydrogen, nitrogen, oxygen, sulphur, etc.) Most of the organic material in Murchison exists as a complex macromolecular substance, meaning a component comprising many hundreds of atoms which has no clearly defined structure. This is analogous in many ways to the material that on Earth occurs in some sedimentary rocks and is responsible for generating oil. Of the organic compounds that are not part of macromolecular structures, many different constituents are present. Some examples of compounds in meteorites are shown in Figure 8.5. Benzene is a hydrocarbon, familiar to many people because of its carcinogenic properties. Adenine and alanine are, respectively, examples of a purine and an amino acid. On Earth, purines are found in nucleic acids and amino acids in proteins. That is not to say that there are nucleic acids or proteins in meteorites! What is present in Murchison are the building blocks of these materials. These same building blocks were presumably brought to the surface of the Earth early in the history of the Solar System, and may have triggered the development of life.

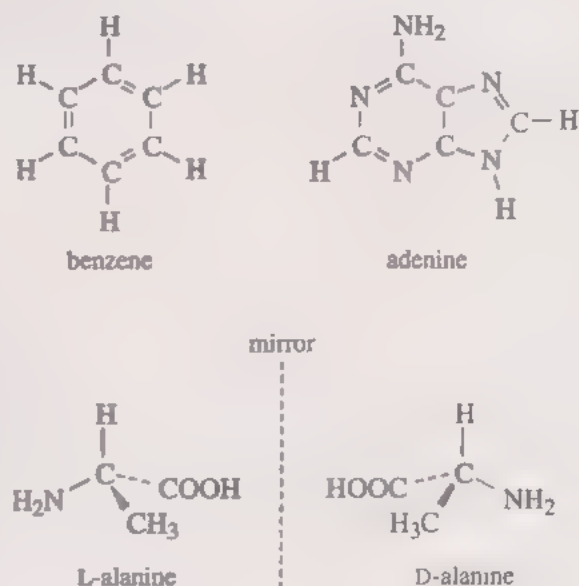


Figure 8.5 Three examples of organic compounds from Murchison. Benzene is a hydrocarbon (a compound made only of hydrogen and carbon). Adenine is a purine, a more complicated molecule than benzene, which includes nitrogen in its structure. Alanine is an amino acid; it can exist in two different structural forms, L and D, which are mirror images of each other. (The dashed lines denote bonds directed behind the plane of the paper, and the wedges are bonds coming up out of the plane of the paper.)

Given that organic compounds abound in the Solar System beyond the Earth (not only in primitive meteorites but also in comets, and probably within the satellites of the giant planets), mechanisms which might lead to the large-scale production of them in the solar nebula are of great interest. It transpires that there are many different ways in which organic materials can be formed. It is not possible to describe all of these here; we will consider just one possibility. If we continue with the notion of a condensation sequence, then at temperatures below 400 K, gases such as CO (carbon monoxide) and H_2 undergo reactions on the surfaces of earlier-formed mineral grains to form organic compounds. This mechanism, along with others proposed, produces organic materials abiotically, i.e. life-forms are *not* involved. Can we be absolutely certain of this? Let us consider the case of amino acids, since these are so closely linked with life. There are many different amino acids, and some of the ones found in Murchison do not form biologically on Earth. At very least, we can be sure that these are not simply terrestrial contaminants. The origin of amino acids is further elucidated from a study of the spatial arrangement of their constituent atoms. Amino acids are known to exist in one of two different structural forms, the L or the D form, which are the same shape but not superimposable, like a left-hand and a right-hand glove, they are mirror images of each other. Thus, as shown in Figure 8.5, the L-alanine and the D-alanine molecule have the same atoms connected together in the same order, but the groups joined to the central carbon atom are arranged differently in space. On Earth, biological activity favours the production and utilization of only the L form of most amino acids. However, L-alanine and D-alanine occur in roughly equal abundances in Murchison, which demonstrates that they did not form biologically.

Based on your experience with the study of refractory inclusions, can you think of a way in which we could try and discern whether all the organic compounds in Murchison had a common origin?

■ One possibility would be to determine the isotope compositions of the constituent elements (carbon, hydrogen and nitrogen) to see if these show any obvious relationships.

The carbon and nitrogen isotope compositions of organic compounds in Murchison are generally fairly uniform, being not too far removed from values found in organic compounds on Earth. The similarity in isotope compositions is taken to imply that the majority of organic compounds found in Murchison formed within the solar nebula. However, the hydrogen in some of the organics in Murchison is found to be highly enriched in the isotope deuterium (D, or 2H). It has been calculated that the magnitude of the observed enrichment cannot be attained in the solar nebula, even under the most extreme conditions imaginable. On the other hand, large deuterium enrichments have been detected in the interstellar medium, using infrared astronomy. The enrichments are found in organic molecules present in dense clouds and are the result of reactions between molecules and ions. It seems, therefore, that some of the organic compounds in Murchison were produced in the interstellar medium. These have survived the processes of heating and mixing in the solar nebula to end up in the organic material found in carbonaceous chondrites.

FIG 8.6 If the majority of organic compounds in Murchison were formed in the solar nebula, as a consequence of condensation, how can we interpret the survival of interstellar molecules in this meteorite?

8.3.4 Silicon carbide – refractory material from beyond the Solar System

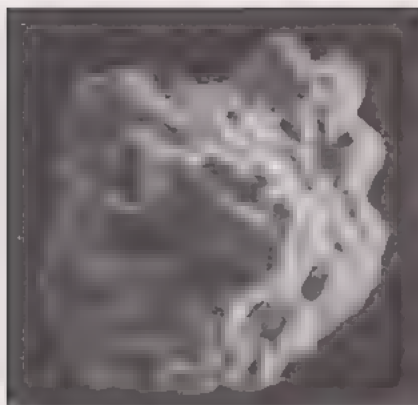


Figure 8.6 Scanning electron microscope image of a silicon carbide grain from Murchison (diameter about $1\ \mu\text{m}$). This mineral condensed around a carbon-rich star over a hundred million years before the Solar System even existed

Nucleosynthesis in red giants, and mass loss from them, are discussed in Book 1, Section 3.4.

Nuclides and nuclear reactions are discussed in *Preparatory science*, Section 2.2

We have seen how Allende and Murchison seem to contain, in addition to materials formed within the solar nebula, remnants of components that pre-date the Solar System. The ^{16}O -enriched minerals in CAIs and deuterium enrichments in organic compounds both represent a ‘memory’ of pre-solar materials. Thus far, it has proved impossible to isolate these pre-solar materials in pure form.

ITQ 8.7 Consider why isolation of these components may be impossible.

In addition to components that show mere glimpses of their pre-solar character, carbonaceous chondrites contain mineral grains that have survived intact the process of Solar System formation and are demonstrably pre-solar in origin. Fortunately, these materials *can* be isolated for study in a virtually pure form. As might be anticipated, the grains in question are extremely robust materials, such as diamond and silicon carbide (SiC). Isotope measurements have confirmed their pre-solar origins. We shall consider only the SiC grains, a typical example of which is shown in Figure 8.6. The grains are generally about $1\ \mu\text{m}$ in diameter and are present at concentrations of a few parts per million in Murchison (but are largely absent in Allende). The ways in which materials of this nature can be studied were documented in the TV programme *Cosmic recycling*.

There are several lines of evidence that help us to identify the likely origins of the SiC grains in Murchison. SiC has been detected spectroscopically in the dust surrounding those red giant stars that are carbon-rich (i.e. have a C/O ratio greater than 1). It can be seen from Figure 8.7 that the carbon isotope compositions of red giants show similarities with data obtained from the SiC grains. Analyses of the xenon contained within the grains shows it to have an isotope composition which was produced by slow neutron capture (s process) nucleosynthesis. A potent site for the operation of the s-process is in stars which have reached the red giant phase of their evolution. We are thus fairly certain that the SiC grains formed as condensates in the stellar winds which outflow from red giants. From these and other considerations it is thought that the stars involved were of relatively low mass (i.e. about $2\text{--}5\ M_{\odot}$).

During their sojourn in the interstellar medium, SiC grains would have been irradiated by high-energy particles in **cosmic rays**. This causes the silicon to transform into a number of nuclides, including a stable isotope of neon, $^{21}_{10}\text{Ne}$; this process is known as **spallation**. By measuring the content of $^{21}_{10}\text{Ne}$, and making some assumptions about the flux of irradiation, it is possible to conclude that the SiC grains are about 140 million years older than the Solar System.

Thus, by studying SiC grains from a meteorite we are able to look back in time to an era that pre-dates the Solar System. We are beginning to comprehend the formation process of the dust particles which were eventually to coalesce and form the solar nebula. It is a sobering thought that the carbon that makes up our own bodies originated in forms like the SiC grains which we can now extract from Murchison.

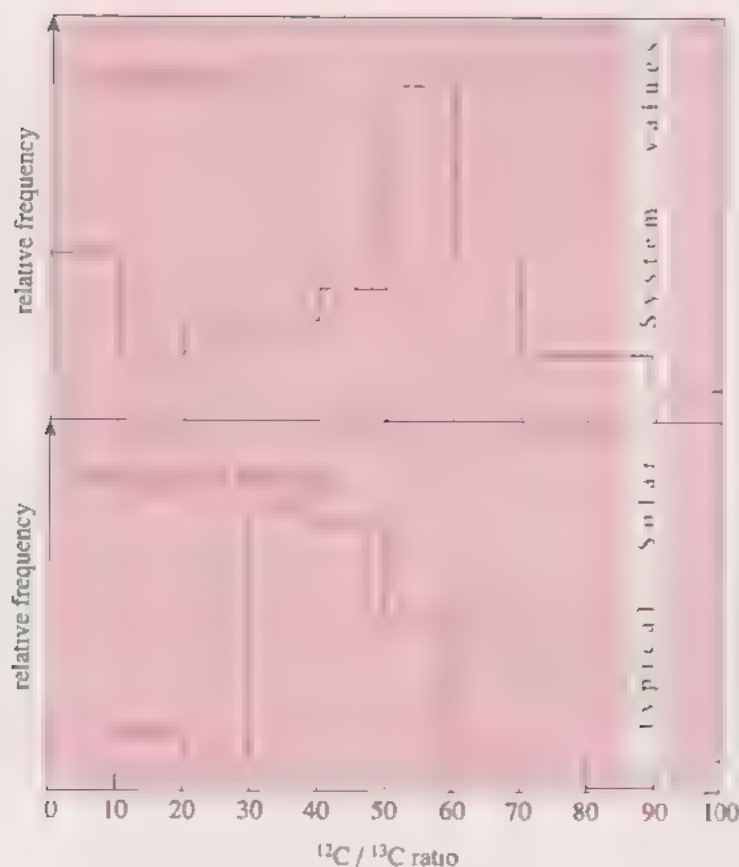


Figure 8.7 $^{12}\text{C}/^{13}\text{C}$ ratios of carbon-rich red giant stars and SiC grains from Murchison. Note the similarities in carbon isotope composition between the two. Note also how different these are to typical Solar System values.

Summary of Section 8.3 and SAQs

- 1 Allende and Murchison are representatives of a chemically primitive group of meteorites known as carbonaceous chondrites. By chemically primitive we mean having a composition which approximates that of the Sun. In detail Murchison is more like the Sun than Allende, although a further group of carbonaceous chondrites, the C1 samples, are the most primitive of all meteorites.
- 2 The carbonaceous chondrites are referred to as cosmic sediments because they represent a collection of materials, formed in different thermal regimes, which settled out from the solar nebula onto the surfaces of parent bodies.
- 3 Refractory inclusions, composed of minerals containing elements such as calcium and aluminium, are present in Allende and Murchison. These materials probably formed early in the solar nebula.
- 4 The oxygen isotope compositions of refractory inclusions show that the solar nebula was never completely homogenized. In other words, these entities retain a 'memory' of the pre-existing materials which became incorporated into the solar nebula.
- 5 Carbonaceous chondrites also contain materials formed at low temperatures, such as carbonates, sulphates and hydrated minerals. The carbonates and sulphates were formed during secondary aqueous activity on the parent bodies.

- 6 Murchison contains 2.5% by mass of carbon in the form of organic compounds. The exact formation mechanism of these materials is unclear but it seems that the majority were formed within the solar nebula, possibly as an extension of the condensation sequence. A small fraction of the organic compounds is demonstrably pre-solar in origin and clearly shows that such material was able to survive the conditions prevailing in the solar nebula.
- 7 Murchison contains grains, in the form of diamond and SiC, that pre-date the solar nebula. The SiC appears to be material formed as condensates in the stellar winds of carbon-rich red giant stars.

SAQ 8.4 (Objective 8.4) Consider the refractory elements Ca and Al shown in Figure 8.3. What differences in chemical composition exist between C1, C2 and C3 meteorites? How could this be explained?

SAQ 8.5 (Objective 8.5) Different sorts of matrix materials from Murchison have oxygen isotope compositions which on a plot of $\delta^{17}\text{O}$ versus $\delta^{18}\text{O}$ fall on line with a slope of about 0.5. This line is parallel to, although displaced from, the TFL. How can the results be interpreted?

SAQ 8.6 (Objectives 8.6 and 8.7) Allende contains carbon mainly in an amorphous form but also as diamonds. How might we try and assess if these elemental forms of carbon are related?

8.4 Juvinas and Cape York – highly processed materials

8.4.1 Melting processes in asteroids

Juvinas is an achondrite which, as you should remember, means that it is a stony meteorite containing no chondrules. It is thought to have been formed near to the surface of its parent body by an igneous process, i.e. a high-temperature melting event which involved the creation of a magma, of which Juvinas is a fragment. Cape York is also the product of melting, although it is from a different parent body to Juvinas. Cape York is an iron meteorite and represents part of an asteroid's core.

Because of melting, Juvinas and Cape York no longer retain their primary chemical and physical characteristics. They are referred to as differentiated meteorites.

Can you remember what differentiation is?

■ It is the process by which a hot, possibly molten, planetary body segregates into compositionally distinct layers of different density (Subsection 3.3.2).

Consideration of iron will help to explain differentiation. When the parent bodies of meteorites were first formed they probably all had a chondritic composition. We know that in primitive chondrites iron constitutes 20–25% by mass and is present in silicates and sulphides, and as the metal itself. However, in Juvinas, iron is present at about 14% by mass, and is almost entirely in silicate minerals – so clearly there is a deficiency of iron in this meteorite. The depletion can be explained by the effects of partial melting (Subsection 5.2.1), whereby the magma that solidified to form Juvinas became separated from its source region, which was left with a relatively higher proportion of iron. The extent of partial melting was fairly limited in this case – thus, although the source region became enriched in iron it probably did not form into a metallic core. In contrast, Cape

York is >90% iron metal, which represents the end-point of total melting. In this case the silicates have been completely removed, leaving a residue of iron–nickel metal.

The observation of meteorites with a variety of iron contents led early workers to propose that all of these materials came from a single disrupted planet which once occupied an orbit between Mars and Jupiter. The irons were thought to represent the core of this planet and the igneous samples were considered to be derived from its crust or mantle. The chondrites were assumed to represent samples of unmelted material which had survived unaltered at the planet's surface. This single-parent model has long since been untenable as detailed chemical and isotope evidence demand many different parent bodies for meteorites. It is now considered that the larger asteroids are planetesimals which were large enough to have undergone a variety of geological-type processes, but (as noted in Subsection 2.5.4) were prevented from accreting into a single planetary body.

8.4.2 Juvinas – a basaltic partial melt

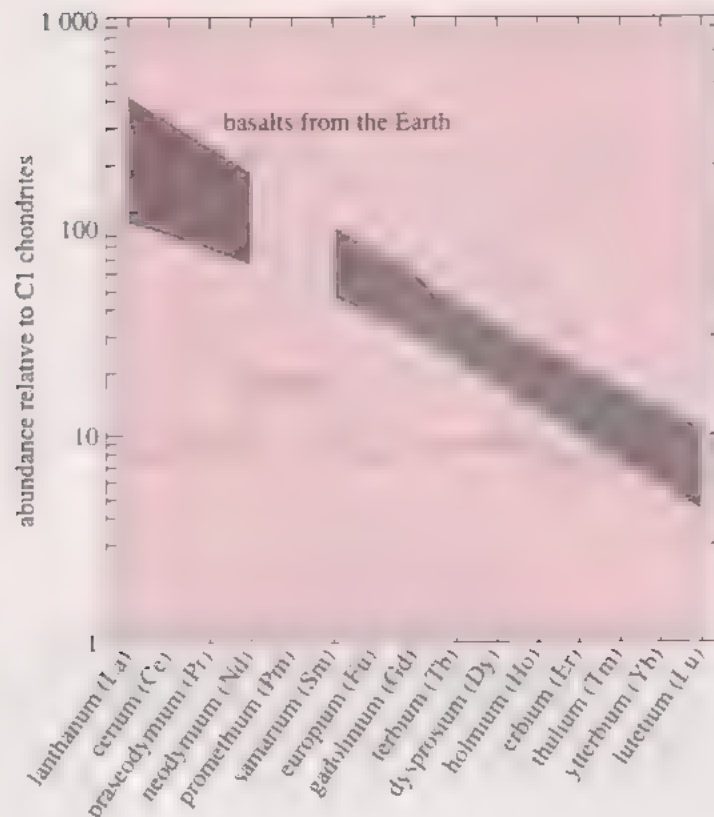
Juvinas is one of a group of about 60 related meteorites. The parental material of Juvinas and its associates must have been extruded onto the surface of an asteroid as a thin layer of magma which then cooled relatively quickly. By studying the chemistry and mineralogy of Juvinas we can learn about the magma itself and also about the source region from which it was extracted.

Looking at a specimen of Juvinas shows it to be a basalt, i.e. comprised mostly of the silicate minerals pyroxene and feldspar. From the mineralogical evidence it is possible to conclude that the sample crystallized within the temperature range 1 300–1 500 K. Be sure that you understand this is magmatic temperature, not a measure of temperature during condensation from the nebula.

As already stated, the heating and melting episode that produced Juvinas has acted to change the original chondritic composition. But how can we be absolutely sure that the parent body of Juvinas had a chondritic composition to start with? All we have to work with is a sample of solidified magma, and we have already seen that the chemical composition of this has been changed by the melting process. Consider two more ways in which the chemistry has been changed. Firstly, you saw in Subsection 8.2.1 that carbonaceous chondrites contain carbon – however, Juvinas contains hardly any carbon. This is because any organic materials that were once present in the parent body have been destroyed by heating and lost as H₂O and CO₂. Secondly, refractory elements, such as osmium, rhenium and iridium, are depleted relative to chondrites – like iron, these were left behind in the source region during partial melting. Fortunately there is a way to assess the original chemical composition of Juvinas. This is by consideration of the **rare-earth elements (REE)**, a group of 15 related elements spanning from atomic number 57 (lanthanum) to 71 (lutetium).

The REE are not present in any meteorite samples at high enough concentrations to form minerals in their own right. Instead they are found as substitutions for other elements in host minerals. The different mechanisms by which REE become incorporated into minerals tend to produce characteristic abundance patterns. The abundances of REE in Juvinas are shown in Figure 8.8, where it can be seen that when compared to C1 chondrites, they are uniformly enriched by a factor of about 8. The reason for the enrichment is that during partial melting the REE have become concentrated in the magma rather than the residual solid. Note that the REE in Juvinas plot on what is roughly a horizontal line. This is known as a chondritic pattern since chondrites also plot on a horizontal line (at a value of 1). In contrast, data from basalts found on Earth plot on a steeply sloping line, which demonstrates that these rocks have formed by a more complex series of processes than the simple melting experienced by meteorites like Juvinas.

Figure 8.8 Rare-earth element abundances in Juvinas and in basalts from the Earth. The data are related to the abundances present in C1 chondrites (which, if plotted, would fall on a horizontal line at a value of 1). Since different rock types show a wide range of REE abundances it is useful to plot the data on a logarithmic scale. Note that there are no data for promethium since this element is entirely radioactive.



Let us now consider how the parent body of Juvinas became heated to temperatures great enough to produce magmas.

Can you think how the parent body may have been heated?

■ As outlined in Subsection 3.5.2, thermal energy derived from accretion, impacts, core formation, radioactive decay and tidal heating, are all potential causes of melting in planetary bodies.

For various reasons, we can discount all but one of these heat-producing mechanisms as causes of melting in asteroids. A small body, such as an asteroid, has a large surface area/volume ratio (compared to a large body) and so thermal energy within the body is rapidly radiated away from its surface. Thus, although kinetic and gravitational energy can be converted to heat during accretion, for small bodies this source is insufficient to cause melting. Heating through asteroid collisions may be a possibility, but if the process is too violent the bodies will be totally vaporized instead. Heating by core formation is discounted since, as you have already seen, a metallic core was not formed in the case of the Juvinas parent body. Radiogenic heating by long-lived radioactive isotopes, such as $^{40}_{19}\text{K}$, $^{232}_{90}\text{Th}$, $^{235}_{92}\text{U}$ and $^{238}_{92}\text{U}$, can produce high temperatures in large planetary bodies, but is again insufficient in small ones. Tidal heating may be an important process for the satellites of the giant planets, but not for asteroids.

So, how *did* asteroids melt? There are two possibilities. The first involves heat production by electromagnetic induction, which you have not met before in this Course. This could arise as the bodies moved through a dense solar wind-like plasma, as might have been produced by the young Sun (during its T Tauri phase). However, there is still some debate about the details of this period of evolution and so it is impossible to assess the significance of this mechanism. A further plausible way of producing heat in asteroids is by short-lived isotopes, such as $^{26}_{13}\text{Al}$ (which you met in Subsection 3.3.2). Since the half-life of $^{26}_{13}\text{Al}$ is only 7.2×10^5 years, the heat potential from this source is confined to the very early stages of Solar System development. The $^{26}\text{Al}/^{27}\text{Al}$ ratio in the interstellar

medium is about 1×10^{-5} (note that $^{27}_{13}\text{Al}$ is the only stable isotope of aluminium). If aluminium with this isotope composition was present in chondritic materials, the heat energy released by the radioactive decay of $^{26}_{13}\text{Al}$ would be sufficient to melt an asteroid of a few kilometres diameter.

ITQ 8.8 $^{26}_{13}\text{Al}$ decays by positron (e^+) emission to give stable $^{26}_{12}\text{Mg}$. How could you tell whether a mineral component from a meteorite had once contained $^{26}_{13}\text{Al}$? In what sorts of minerals would the identification of previous $^{26}_{13}\text{Al}$ radioactivity be most easily recognized?

Unfortunately, as far as Juvinas is concerned, it is impossible to discern whether $^{26}_{13}\text{Al}$ was responsible for melting. However, we do know that $^{26}_{13}\text{Al}$ was present in the solar nebula since evidence of its decay product $^{26}_{12}\text{Mg}$ has been found in CAIs from meteorites like Allende and Murchison. That the chondritic meteorites did not melt shows that, in those samples at least, the overall concentration of $^{26}_{13}\text{Al}$ was not high enough. Thus, while it is a commonly held belief that $^{26}_{13}\text{Al}$ was a heat source in the early Solar System, the exact influence is still being actively researched.

Notwithstanding the lack of a consensus regarding the heat source for asteroidal melting processes, the timing of this event can be reliably established by using a radiometric dating technique (Box 3.1). This involves using a mass spectrometer to measure the quantity of a stable isotope which has formed from the decay of a radioactive isotope. A good example here is the Rb–Sr system, in which the amount of stable $^{87}_{38}\text{Sr}$ produced by the decay of a given amount of $^{87}_{37}\text{Rb}$ is a measure of the formation time of a sample, i.e. the time at which it solidified, or crystallized – this being the moment when the strontium began to be trapped in the same mineral grains as the rubidium that gave it birth. The Rb–Sr age of Juvinas is 4.50 ± 0.07 billion years, whereas Rb–Sr ages for a range of other meteorite types span 4.45–4.53 billion years. (There are samples with ages younger than this range, but these document later impact events rather than the initial (partial) melting of the asteroid.)

If the ages obtained from the oldest meteorites represent the age of the Solar System, what constraints do the Rb–Sr ages of different types of meteorites put on formation conditions in the solar nebula?

■ Melting processes on meteoritic parent bodies took place only a short time after the Solar System was formed.

The photometric properties of meteorites like Juvinas resemble those of Vesta, one of the largest asteroids (Table 1.1c). This large body (250 km radius) is still demonstrably intact. If Juvinas came from Vesta then we can deduce that it must have been ejected from the asteroid's surface. In other words, we cannot expect to have samples of its interior. However, we do have several hundred specimens of iron meteorites which are thought to represent the cores of asteroids. Cape York is one such sample.

8.4.3 *Cape York – a sample from the core of an asteroid*

Cape York is part of one of the largest meteorite showers from which we have samples. The meteorite fell to Earth over 1 000 years ago in Greenland; eight individual specimens have so far been collected, totalling 5.9×10^4 kg. In historical times, fragments of the meteorite were used extensively by the Inuit (Eskimos) in the manufacture of tools. When tools were traded between the Inuit and explorers, who visited the island in 1818, it became apparent to the scientific community at large that there must be an iron meteorite somewhere in Greenland. In 1897 the largest known fragment of the meteorite (3.1×10^4 kg)

was found and taken by sea to New York; another large mass (2.1×10^4 kg), which was unknown to the Inuit, was removed in 1965 by a Danish expedition.

Iron meteorites originate from asteroidal bodies – the processes which contributed heat were similar to those which produced meteorites like Juvinas. However, iron meteorites are not the result of partial melting. They represent the end-product of differentiation and segregation of metal. In other words, the parental magmas of iron meteorites were, at some point in their history, completely molten. During this process, low-density silicates became removed by essentially floating upwards towards the outermost part of the asteroid. As such, most iron meteorites represent samples of the cores of asteroids. We can be certain that a number of different parent bodies are represented among all the iron meteorites, since the concentrations of trace elements, such as iridium, gallium and germanium, vary in ways which could not be produced in a single body.

TIQ 8.9 List the mechanisms that are likely to have contributed heat to asteroids.

Iron meteorites are alloys of iron and a relatively small proportion of nickel. Cape York, for instance, contains about 7.5% nickel. When a meteorite like Cape York is sawn open, polished and then etched with a mild acid, a characteristic pattern can be observed. This is known as the **Widmanstätten pattern** and while it can be reproduced to a very limited extent in the laboratory it is not a normal feature of terrestrially produced iron. An example of an iron meteorite displaying the Widmanstätten pattern can be seen in Plate 2.47a. The pattern is caused by the segregation, during extremely slow cooling, of nickel-rich and nickel-poor alloys. Cooling rates, calculated from the different scales of pattern that have been found in iron meteorites, vary from 0.5–500 K per million years, differences which reflect parent bodies of different sizes – larger bodies cooled more slowly. Some iron meteorites, which have related chemical compositions but different cooling rates, are considered to have formed at varying depths within an individual body, i.e. samples closer to the centre cooled more slowly. If this is true, the parent asteroids cannot have had single cores of iron, which has a high thermal conductivity, but must have been composed of individual pods of metal surrounded by silicates. Such parent bodies are said to have a *raisin-bread* structure.

Most iron meteorites are not composed of pure iron–nickel metal. They contain various quantities of phosphorous containing minerals, sulphide nodules and graphite. Other irons contain inclusions of silicates, which have chemical compositions similar to chondrites.

Does the presence of silicate inclusions pose any constraint on the formation conditions of iron meteorites?

■ The presence of silicate inclusions shows that these samples were never completely molten.

Thus not all iron meteorites were formed by processes that produced total melting. Those with silicate inclusions are considered to be the result of impact melting and mechanical mixing of iron-containing materials with silicate-rich bodies. This happened early in the history of the Solar System when newly-formed bodies were colliding with each other. Thus, not all iron meteorites can be thought of as being the cores of asteroids.

- 1 Juvinas and Cape York are examples of meteorites which have been subjected to melting processes on asteroidal bodies. We refer to these meteorites and their parent bodies as differentiated.
- 2 Meteorites like Juvinas are comprised of materials which are the result of primary igneous activity, i.e. magmas produced via the partial melting of a chondritic source. They may all come from a single asteroid (Vesta).
- 3 Most iron meteorites are the result of the total melting which produced layering within asteroids. In the case of Cape York, the iron–nickel body formed the core of an asteroid. However, iron meteorites could also arise from asteroids which have a raisin-bread structure, in which there are discrete pods of iron–nickel metal.
- 4 Some iron meteorites were not formed by processes that produced total melting but were instead produced by impact melting and mixing at the surfaces of asteroidal bodies.
- 6 The short-lived radioactive isotope $^{26}_{13}\text{Al}$ may have been responsible for heating and melting asteroids.
- 7 The formation ages of meteorites can be assessed by radiometric dating (e.g. using the Rb–Sr system).

SAQ 8.7 (Objective 8.2) Some of the observable asteroids, classified as type M, have a metallic composition. What does this tell us about processes within the asteroid belt?

SAQ 8.8 (Objectives 8.8 and 8.9) Kodaikanal is an unusual iron meteorite which has a Rb–Sr age of 3.8 billion years. This is within the range recorded in the majority of impact-produced rocks from the Moon (3.8–4.0 billion years). What can we infer here?

8.5 EET A79001 – a meteorite from Mars

8.5.1 Collection in Antarctica

EET A79001 is a meteorite collected from Antarctica – hence the lettering and numbering system. ‘EET’ refers to the collection locale (which, in this instance, was Elephant Moraine), ‘A’ designates the collection trip and ‘79’ is the year of collection (1979). The identifying number, 001, signifies that it was the first meteorite to be classified upon return of the samples to the curatorial facility. When the collectors, who are often meteorite researchers, spot rare or otherwise unusual samples in Antarctica, a note is made to give them priority treatment during the preliminary classification procedure. EET A79001 was one such sample, and on this particular occasion the decision to promote its investigation could not have been more justified, since the meteorite is now widely believed to come from Mars.

8.5.2 Why do we think that EET A79001 comes from Mars?

EET A79001 is a so-called **shergottite** (one of only five samples of its type). A photograph of EET A79001 is shown in Plate 2.50. For various reasons shergottites are linked with two further categories of meteorites known as **nakhlites** (pronounced ‘nahk-lights’: three individuals) and **chassignites**

(pronounced 'sha-sig-nights'; only one sample). These nine samples, and one other as yet unidentified sample, collectively are referred to as the **SNC meteorites** (where S, N and C denote shergottites, nakhlites and chassignites). All of the SNC meteorites are igneous rocks formed by crystallization from a magma. In appearance the shergottites are similar to meteorites like Juvinas, being medium-grained rocks of basaltic composition. However, one important fact distinguishes SNC meteorites – they have relatively young formation ages.

The crystallization ages of SNC meteorites range from 0.2–1.3 billion years. What are the implications of this?

■ These meteorites were formed late in the history of the Solar System. Wherever they were formed, at least some part of their parent body had to have been melted 0.2 billion years ago (or remained molten until this time).

As stated in Subsection 8.4.2 the ages of most meteorites cluster around 4.5 billion years and represent samples which formed on parent bodies of asteroidal size (i.e. a few hundred kilometres in diameter). Remember that asteroids were heated and cooled relatively early in the history of the Solar System. Some meteorites have younger ages than 4.5 billion years, but these are the result of later impact melting which acts to re-set the radiometric dating systems. It is apparent from the textures of SNC meteorites that they are not impact-produced melts. The only reasonable environment that retains sufficient heat to produce melting 0.2 billion years ago is a parent body of planetary dimensions, so there can only be a few candidates for the source of SNC meteorites – i.e. Mercury, Venus, Earth, the Moon, Mars, or Io.

For a number of reasons it is considered that Mercury is extremely unlikely to be the SNC parent body. Mercury is probably a volatile-poor planet, unlike the SNC meteorites. Furthermore, being so close to the Sun, it is unlikely that material ejected from Mercury could be transported to the orbit of Earth. Venus has a dense atmosphere (Table 1.1a) and a large gravitational field (about the same as Earth) – removal of material from this planet would require velocities so high that frictional heating in the atmosphere would cause severe melting (or even complete vaporization in the case of small pieces of rock). The Earth is excluded because the oxygen isotope compositions of SNC meteorite samples do not fall on the TFL (see Subsection 8.3.2). The Moon is also discounted on the basis of oxygen isotope composition, and because lunar volcanism ceased about 3.2 billion years ago.

Mars would seem to be a good possibility for the source of the SNC meteorites, if only by default. In support of this it should be noted that Mars has a tenuous atmosphere (surface pressure currently 6 mbar) and the gravitational field is comparatively low (less than half that of the Earth) – thus, solid materials could be ejected without vaporization – and it is a lot closer than Io.

Can you think of ways in which the SNC meteorites could be removed from the surface of the planet?

■ Ejection could be caused by a volcanic eruption or an impact.

Volcanic processes have been considered but rejected on various grounds. The most plausible mechanism involves removal following an impact onto the Martian surface by a meteoroid, or comet. Early attempts to model this process were inconclusive. There was once a persuasive argument that SNC meteorites could not come from Mars because no lunar meteorites had ever been found on Earth. From dynamical considerations it should be easier to eject samples from the Moon and for them to arrive on Earth as meteorites, than it would be in the case of Mars. There the matter rested until the discovery of ALH A81005 (another Antarctic meteorite). There is universal agreement that this sample has a lunar origin. (Several more lunar meteorites have subsequently been recognized.) We can be sure that these meteorites come from the Moon because we can make detailed and direct comparisons with lunar samples returned to Earth by space missions.

Acceptance of a lunar origin for some meteorites has removed doubts about the ability of projectiles to eject materials from the surfaces of planetary bodies.

11Q 8.10 Based on your knowledge of the surface conditions of Mars and the Moon, how might you expect Martian and lunar meteorites to differ?

We have already stated that SNC meteorites are not impact melts but are the result of igneous activity, analogous in many ways to rocks formed at or near the Earth's surface. However, SNC meteorites were subjected to an impact event which was violent enough to accelerate them to a speed greater than the escape speed of Mars (5 km s^{-1}). Intuitively therefore, it may be expected that the meteorites would show some evidence of this process. Indeed, shergottites and chassignites record the effects of shock, though the nakhlites are unshocked (which poses a constraint on theoretical modelling of the impact event). A schematic representation of calculations pertaining to large impact craters on Mars is shown in Figure 8.9. You should recognize the general features of the process, familiar to you from Subsection 4.2.1 and video sequence 6, but note that in this model the incoming projectile, in addition to pulverizing the target rocks and producing a crater, also causes a very thin layer of ejecta to be propelled away from the impact site. Some of these pieces of ejecta, which are unmelted and unshocked, can escape from the planet's surface. Other escaping ejecta components could be melted or shocked.

In contrast to the nakhlites, EET A79001 has been subjected to shock pressures of $8 \times 10^{10} \text{ Pa}$. This has resulted in the transformation of plagioclase feldspar minerals into a glass, known as maskelynite. In addition, small amounts of olivine and pyroxene crystals have been fractured and in some cases melted to give a shock-produced glass, which has formed into veins. Analyses of gases trapped in the shock-produced glass from EET A79001 shows them to be chemically and isotopically distinctive. As you will see shortly, they bear many similarities to gases in the Martian atmosphere. This provides the most compelling reason to pronounce that Mars is the parent body of the SNC meteorites.

Maskelynite is individual crystals turned to glass by intense pressure (because of shock waves); it is *not* the same thing as glass produced by impact melting

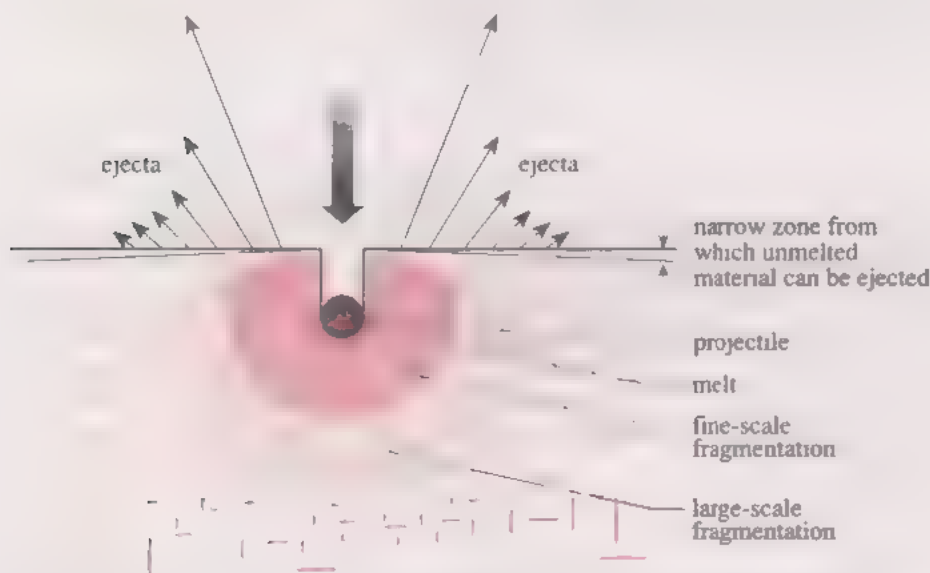


Figure 8.9 Schematic diagram of a crater-forming event on Mars. No physical parameters are given since this is a generalized case (although imagine that it pertains to projectiles of kilometre size travelling at around 10 km s^{-1}). Note that the impact produces melting and intense fragmentation in the target rocks, with fragment size increasing away from the impact site. However, in addition, unshocked and unmelted materials can be ejected from a very thin surface layer. Close to the impact site, these may be travelling sufficiently fast to overcome the escape speed of Mars and are, as a consequence, ejected from the planet.

8.5.3 Nitrogen on Mars

The Viking spacecraft visited Mars in 1976 (Section 6.2). During these missions the Viking landers determined the constitution of the Martian atmosphere (Table 8.1). The abundances of some of these gases are given in Figure 8.10 where they are compared with the abundances of the same gases contained within the shock-produced glass from EET A79001. The similarities are self-evident.

During the impact event that is thought to have removed EET A79001 from Mars, localized melting occurred within the sample. These melts cooled extremely rapidly to form a glass containing trapped atmospheric gases. Laboratory experiments have reproduced this phenomenon and shown that the isotope compositions of the trapped gases remain the same as those of the ambient atmosphere. This is an important observation which implies EET A79001 could contain an unadulterated sample of the Martian atmosphere.

Table 8.1 The chemical composition of the Martian atmosphere (data determined by the Viking landers)

Gas	Formula	Volume ratio ^a
Carbon dioxide	CO ₂	0.953
Nitrogen	N ₂	2.7×10^{-2}
Argon	Ar	1.6×10^{-2}
Oxygen	O ₂	1.3×10^{-3}
Carbon monoxide	CO	7×10^{-4}
Water vapour	H ₂ O	3×10^{-4}
Neon	Ne	2.5×10^{-6}
Krypton	Kr	3×10^{-7}
Xenon	Xe	8×10^{-8}
Ozone	O ₃	3×10^{-8}

^a This is the fraction by *number* of the atoms or molecules present.

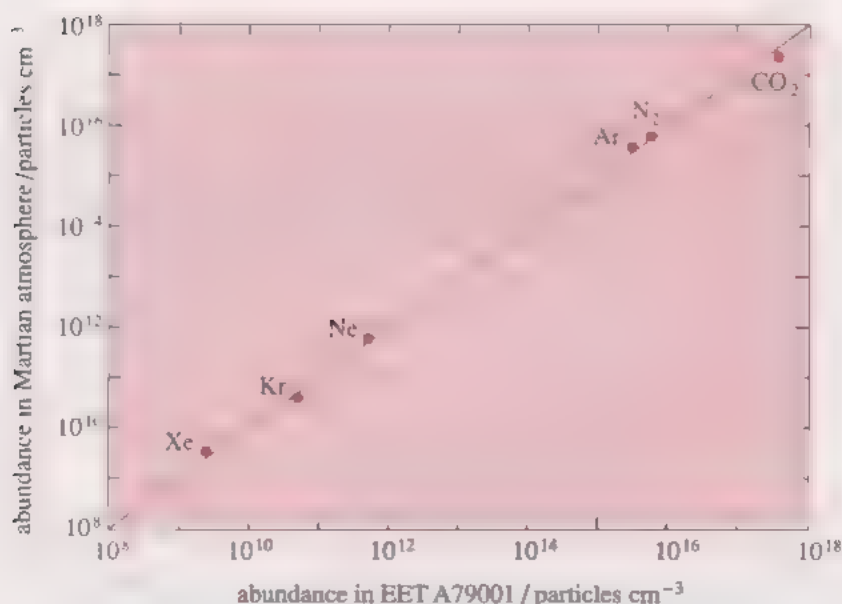
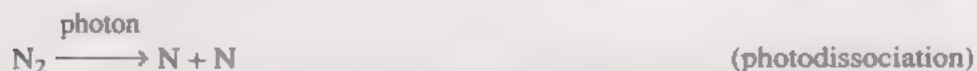
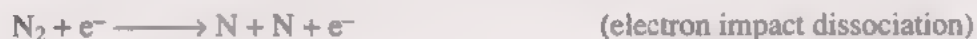
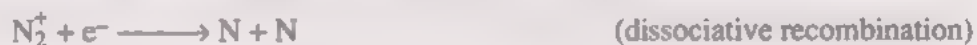


Figure 8.10 Plot of the abundances of gases in the Martian atmosphere (measured by Viking landers) versus the abundances in the glass from EET A79001. Both axes are logarithmic. Note the extremely good correlation, which forms part of the evidence supporting a Martian origin for SNC meteorites.

The isotope compositions of some of the gases in the Martian atmosphere were determined by the Viking landers. These data, along with the equivalent values for the Earth's atmosphere, are given in Table 8.2. The errors on the terrestrial data are relatively minor and can be effectively ignored – in contrast, there are large uncertainties associated with the measurements made during the space missions. This is the inevitable consequence of attempting to make measurements using remotely operated instrumentation. Nevertheless, some of the Earth–Mars differences are significant. Also, by analysing gases in EET A79001 it should be possible to reduce the analytical uncertainties associated with isotope ratio measurements. We will consider the case of nitrogen, but firstly it will be necessary to consider why there might be such a difference in the $^{14}\text{N}/^{15}\text{N}$ ratios between the Earth and Mars. Compared to the Earth, the Martian atmosphere is highly enriched in ^{15}N ($^{14}\text{N}/^{15}\text{N} = 165$ for Mars as compared to 272 for the Earth).

Let us for the moment assume that, at the time of planetary formation, the Earth and Mars were endowed with nitrogen of a similar isotope composition. It is known that all of the nitrogen originally present on Earth remains on the planet, mostly in the atmosphere. Thus, we can conclude that the nitrogen in the present-day atmosphere on Earth has the same isotope composition as it did to start with. In contrast, during the history of Mars, nitrogen was irreversibly lost from the atmosphere to outer space – this has ultimately resulted in a change in the isotope composition of the gas which remains on the planet. The gas has been lost in the form of nitrogen atoms, which may be formed in the atmosphere by three different processes as follows:



These processes can impart enough energy into the nitrogen atoms to give them speeds in excess of the escape speed of Mars. Although the same processes take place in the Earth's atmosphere, the escape speed in this case is much higher and so the atoms are retained. In the case of Mars, both ^{14}N and ^{15}N atoms are lost, but at different rates – ^{14}N (having lower mass) is lost more readily. The net result is that the present-day Martian atmosphere has become enriched in ^{15}N . From the magnitude of the nitrogen isotope fractionation it is possible to calculate that there was originally between 5 and 150 times the present-day quantity of nitrogen in the Martian atmosphere.

When we measure the isotope composition of nitrogen in the glass fraction of EET A79001, we find that the $^{14}\text{N}/^{15}\text{N}$ ratio is at most 209 (compare Table 8.2).

Can you now think of a way in which the nitrogen isotope composition of the gases in EET A79001 could be explained?

■ It could be interpreted as a historical sample of the Martian atmosphere, representing nitrogen from an intermediate stage of evolution which was trapped some time in the past.

Before we can address this issue we will consider further the age of EET A79001. We know that EET A79001 crystallized from a melt 0.2 billion years ago. However, the glass was formed some time later when the meteorite was ejected from the Martian surface. To derive the date of ejection involves measurement of what is known as the **cosmic ray exposure age**. After ejection, while the sample was in space, certain isotopes would have been produced inside the meteorite as a result of cosmic ray interaction. This is analogous to the build-

Table 8.2 Isotope ratios in the atmospheres of the Earth and Mars (data for Mars obtained by the Viking landers)

Ratio	Earth	Mars
$^{12}\text{C}/^{13}\text{C}$	90	90 ± 9
$^{16}\text{O}/^{18}\text{O}$	488	500 ± 50
$^{14}\text{N}/^{15}\text{N}$	272	165 ± 16.5
$^{40}\text{Ar}/^{36}\text{Ar}$	296	$3\,000 \pm 300$
$^{129}\text{Xe}/^{132}\text{Xe}$	0.98	2.5 ± 0.25

The analytical techniques used to measure the nitrogen isotope ratios rule out terrestrial contamination.

Relative to the time spent in space by SNC meteorites, the time spent on the Earth's surface before collection and analysis is negligible.

up of $^{21}_{10}\text{Ne}$ in SiC, discussed in Subsection 8.3.4. By analysing the quantities of cosmic ray-produced species in EET A79001 we can determine how long the sample was in space. It transpires that it was ejected from Mars 0.6 million years ago, long after the meteorite material first formed.

Shergottites, other than EET A79001, have exposure ages of about 2.5 million years. The nakhlites and chassignites, on the other hand, have exposure ages of 11 million years. What is the implication of these different ages?

■ The SNC meteorites were not all ejected from Mars at the same time. Rather, separate events are necessary to explain the exposure age data. It would seem that ejection of material from Mars, and subsequent capture of this debris by the Earth, was not a unique event.

As the ejection of EET A79001 from Mars took place relatively late in the history of Mars, the sample of the atmosphere trapped in the glass is only 0.6 million years old. It is unlikely that the nitrogen isotope composition of the Martian atmosphere has changed significantly between 0.6 million years ago and the present day. Thus, the $^{14}\text{N}/^{15}\text{N}$ value of 209 for the nitrogen trapped in EET A79001 cannot, after all, be explained as a half-way stage in the atmospheric evolutionary process.

It becomes easier to explain the nitrogen isotope composition of the trapped gas in EET A79001 when we take into account the $^{14}\text{N}/^{15}\text{N}$ ratios of other SNC meteorites. These all show values which are like that observed in the terrestrial atmosphere. In the case of the nakhlites, which are unshocked and do not contain any glass or samples of the Martian atmosphere, the nitrogen present is probably from the Martian mantle (this is analogous to nitrogen in terrestrial basalts, which arises from mantle sources). Thus, we can conclude that the nitrogen isotope composition of the Martian mantle and crust is Earth-like. The nitrogen in EET A79001 presumably represents a mixture of atmospheric and crustal sources, which would give a measured $^{14}\text{N}/^{15}\text{N}$ intermediate between the two. It would seem that during formation of the glass in EET A79001, nitrogen from the Martian atmosphere became entrapped along with crustal nitrogen in the target minerals.

By analysing SNC meteorites we have ascertained that Earth and Mars originally had similar nitrogen isotope composition. This original composition, which was not measured by the Viking landers, is important since modelling of Martian atmospheric evolution requires a knowledge of the original nitrogen isotope composition of the planet.

11Q811 On the face of it, the carbon isotope composition of the atmospheres of Earth and Mars are the same and so loss of CO_2 from the Martian atmosphere appears unlikely. What measurements could we make on SNC meteorites to check this?

8.5.1 *Evolution of the Martian atmosphere – the role of water*

As you saw from Figure 8.10, the glass in EET A79001 contains a variety of gases. Through analyses of the noble (inert) gases, a theoretical model has been devised to describe the early evolution of the Martian atmosphere. Before considering this further it should be noted that the isotope composition of xenon in the EET A79001 glass shows that the Martian atmosphere includes gases attributable to carbonaceous chondrites (this can also be inferred from an assessment of Viking data, or by looking in detail at the chemical compositions of the SNC meteorites). It is considered that Mars received its presently observed complement of volatiles from an influx of carbonaceous chondrite-like material near the end of its formation. The early Martian atmosphere was subsequently formed by outgassing from this veneer of chondritic material.

□ As you should remember, carbonaceous chondrites contain hydrated minerals. What will happen to those components during impact processes?

■ They will be heated, thereby liberating water. Continued bombardment will produce a dense atmosphere of H_2O .

Water on Mars is a particularly interesting subject. Observations of the Martian surface show considerable evidence for the action of fluid flow (for instance, see Figures 4.19 and 6.31, and Plates 2.14 and 2.15). However, liquid water is demonstrably no longer present at the surface of the planet.

In Chapter 6 you learned about the evolution of H_2O on planetary bodies. Can you remember what happened to the water on Mars?

■ Some of it now resides in the polar caps and permafrost regions. However, large amounts were also lost over time to outer space.

Two episodes of water loss from the atmosphere to space can be identified. At the present time, photodissociation of water vapour results in the loss of hydrogen and oxygen. This mechanism has been operating on Mars since its formation. In earlier times, because the ultraviolet flux from the Sun was comparatively high, resulting in enhanced levels of photodissociation, the process was more efficient. Even so, the rate is too slow to account for all the water assumed to be lost. Quite clearly a further process was responsible.

Over the first hundred million years of Martian evolution, water was readily converted to hydrogen by reaction with, for instance, iron–nickel metal. The vast amounts of hydrogen produced in this way were very quickly lost from the planet as a rapidly moving flow of gas. During this time, gases heavier than hydrogen were swept away by the rapid flow, and so were also lost from the planet – a process known as **hydrodynamic escape**. The details of this are complicated but its imprint can now be observed in the abundances of noble gases and the isotope composition of xenon in the glass of EET A79001.

With the knowledge that hydrogen has been lost from the Martian atmosphere, how would you anticipate its isotope composition has evolved with time?

■ By analogy with the case of nitrogen, it would be expected that the lighter isotope of hydrogen would be preferentially lost from the atmosphere, resulting in an increase in the D/H ratio with time.

Measurements of water vapour in the Martian atmosphere show a D/H ratio which is about 5 times that of water in the Earth's oceans. (Unfortunately the D/H ratio of water in the glass from EET A79001 has not yet been measured.) On the other hand, Table 8.2 shows that Martian atmospheric water vapour is not enriched in the heavier isotopes of oxygen. Thus, even though oxygen may have been lost from the atmosphere, its isotope composition has remained constant with time. This suggests there is a feedback mechanism between the atmosphere and the surface materials on Mars, whereby any changes in oxygen isotope composition can become re-set. The surface materials include silicates and polar ice, which are far more abundant than the atmospheric constituents. Since water condenses and evaporates on a seasonal basis, any changes in the oxygen isotope composition of atmospheric water are eliminated by isotope exchange reactions taking place at the surface. Since hydrogen constitutes only a relatively minor part of the Martian crust, the isotope composition of this element cannot become re-set. Instead, the non-atmospheric surface materials will gradually become more like the atmospheric component. Consequently we would expect any hydrated minerals present at the Martian surface to be enriched in deuterium, just like the atmospheric water vapour.

We can test this notion by studying the hydrated minerals found in SNC meteorites (from various lines of evidence, most of these minerals appear to be low-temperature weathering components formed at the surface of Mars). When we measure the hydrogen and oxygen isotope compositions we find some surprising results. Firstly, the D/H ratio is not as extreme as expected. Secondly, oxygen measurements do not fall on a line of slope 0.5 on a plot of $\delta^{17}\text{O}$ versus $\delta^{18}\text{O}$ – the oxygen in the hydrated minerals cannot be formed from the silicate components. It would appear that the water present in the hydrated minerals of SNC meteorites is not even Martian! Perhaps a component of the Martian surface is extraneous to the planet – for instance, a long time after Mars formed, water-bearing components may have been delivered in the form of meteoroids or comets. Clearly we still have a lot to learn about the evolution of planets like Mars.

Summary of Section 8.5 and SAQs

- 1 EET A79001 is a basaltic rock, formed by an igneous process at the surface of Mars.
- 2 We have, on Earth, 10 samples of Martian origin; these are referred to as SNC meteorites.
- 3 SNC meteorites span a range of crystallization ages from 0.2–1.3 billion years. Ages within this range are not observed in any other meteorite specimens.
- 4 SNC meteorites were removed from the surface of Mars by an impact. It appears there were a number of ejection events.
- 5 EET A79001 is comprised in part of a shock-produced glass which contains a sample of the Martian atmosphere.
- 6 Nitrogen in the Martian atmosphere is highly enriched in ^{15}N as a result of atmospheric loss processes.
- 7 A large amount of water was once present on Mars but this has been lost from the planet.
- 8 The hydrated minerals in SNC meteorites were formed by low-temperature weathering processes at the surface of Mars. Isotope measurements of hydrogen and oxygen in these minerals seems to suggest an origin that involved a cometary or meteoroidal input.

SAQ 8.9 (Objectives 8.5 and 8.9) What diagnostic features would tell you that an SNC meteorite was not simply a terrestrial rock?

SAQ 8.10 (Objective 8.10) A new basaltic meteorite has just been recovered. Explain how you would assess whether it was an SNC meteorite.

SAQ 8.11 (Objective 8.11) It is possible that Martian soil contains nitrates formed from chemical reactions involving the atmosphere. What $^{14}\text{N}/^{15}\text{N}$ ratios would we expect nitrates to have if they were formed (a) now and (b) at some point in the past?

8.6 Micrometeorites – samples of comets

8.6.1 The Giotto space mission to Halley

The most recent apparition of comet Halley unfortunately gave a rather unremarkable visual display in the night sky (the previous appearance was much more spectacular – Figure 1.9). However, much useful scientific information was gained through measurements made by a variety of ground-based instrumentation. In addition, Halley was visited by a total of five spacecraft which carried out a number of investigations. The most varied set of experiments was carried aboard Giotto, a spacecraft built and operated by the European Space Agency (see table near the end of *Images of the Cosmos*). On 14 March 1986, when at a distance of 1.35×10^{10} km from the Sun, Giotto passed within 600 km of Halley's nucleus. The spacecraft made a series of measurements of the magnetic field and energetic particles, analysed the chemical and isotope compositions of the gas/dust, and captured some spectacular images, such as Plate 2.45.

The breadth of the results obtained by Giotto is too large to describe in detail here – we will concentrate on measurements pertinent to the laboratory study of meteorites and micrometeorites.

| Can you recall, in general terms, the constitution of a comet?

■ In Subsection 1.3.2 you learned that comets are considered to be mixtures of ice and dust. It is the non-volatile dust grains that are of interest for comparison with meteorite samples.

Mass spectrometers aboard the spacecraft that visited Halley were used to analyse the chemical and isotope compositions of ejected dust particles. Some of these particles were found to be silicates with chemical compositions similar to silicate particles in carbonaceous chondrites, while others were interpreted to be individual mineral grains of different composition. By far the most abundant particles, however, were found to be composed of C, H, N and O and are interpreted as being organic-rich grains.

The total dust flux from Halley was determined by a British-led experiment on board Giotto. As the spacecraft approached the comet, the dust flux was seen to increase. At the time of encounter, Halley was only four weeks after perihelion and was consequently still very active. It was found that the comet was losing dust at a rate of about $3 \times 10^3 \text{ kg s}^{-1}$. The dust flux measured by Giotto at closest approach is given in Figure 8.11 – by far the greatest number of particles leaving the comet are relatively small in size (mass $< 10^{-18} \text{ kg}$). However, the total amount of mass lost by the small particles is small compared to the mass lost through the ejection of large particles.

8.6.2 Interplanetary dust

When a comet passes close to the Sun it is heated and so some of the icy components become vaporized. At the same time dust grains, ejected from the surface, enter the interplanetary medium as micrometeoroids. However, this is not the only source of dust in the interplanetary medium. Collisions between asteroidal bodies also produce dust-sized debris. Collectively, this **interplanetary dust** is responsible for faint luminosities in the night sky, known as zodiacal light and gegenschein. These phenomena are due to the scattering of sunlight by dust particles. Let us estimate the rate at which the dust is being created.

By studying the frequency of microscopic impact pits on the surfaces of rocks returned from the Moon, it has been established that the flux of micrometeoroids has remained constant for a large part of the history of the Solar

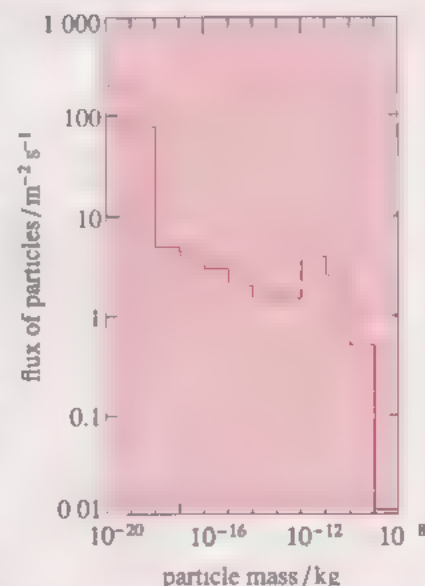


Figure 8.11 The dust flux from Halley versus the mass of the particles. A surprise of this experiment was the relatively large numbers of particles with masses less than 10^{-18} kg .

System. However, it has been calculated that individual grains of interplanetary dust have a lifetime of only about 10^4 years.

What processes will act to destroy, or remove, dust from the interplanetary medium?

■ As alluded to above, dust is constantly intercepted by the larger bodies in the Solar System (for example, dust is captured by the Earth).

Of more significance are losses due to the process whereby particles interact with sunlight. The smallest grains ($< 1\ \mu\text{m}$ diameter) are removed from the Solar System by the effects of a phenomenon known as radiation pressure. Larger particles, on the other hand, lose momentum as they absorb sunlight and then re-radiate it. The inevitable consequence is that they slow down and subsequently spiral in towards the Sun by gravitational attraction. Particle-particle collisions constantly fragment the dust into ever finer particles – since small particles are lost more rapidly than large ones, this helps to accelerate dust loss from the interplanetary medium. The outcome of these loss processes is that for the interplanetary dust cloud to remain roughly intact over long periods, it must be constantly replenished. Calculations show that about $8 \times 10^3\ \text{kg s}^{-1}$ of material are needed to maintain the present cloud.

8.6.3 Laboratory analyses of cometary dust

It can be seen from Figure 8.11 that for the dust emanating from Halley, in addition to there being a relatively higher incidence of very small particles, there is also an enhancement of grains in the range 10^{-11} – $10^{-12}\ \text{kg}$. This corresponds to particle sizes of about 10 – $1\ \mu\text{m}$. Imagine a carbonaceous chondrite, like Murchison, disaggregated into fragments of this size. CAIs and chondrules would fracture into refractory particles of chondritic composition. In addition, the matrix would fragment into particles of hydrated minerals and organic-rich components. Thus, it would seem that there could be a relationship between Halley dust and meteorites. But, is this likely? As you should remember from Section 8.2, most meteorites are thought to come from asteroids, which are ostensibly a different source to comets. However, it has been suggested that some asteroids could be the refractory remnants of comets that have lost all their volatiles after many perihelion passages. Thus, we may expect similarities between comets and certain meteorites.

A more obvious choice of material for comparison with Halley dust would be micrometeorites that can be collected from the Earth's upper atmosphere by high-flying aircraft. These particles are typically $10\ \mu\text{m}$ in size, have roughly chondritic chemical compositions, and densities that range from 0.7 – $2.2 \times 10^3\ \text{kg m}^{-3}$. Although a proportion of them may be asteroidal debris, the majority almost certainly come from comets. Figure 8.12 shows two examples of micrometeorites. One is a highly porous aggregate of small grains of olivine, pyroxene and carbonaceous matter. The other is less porous and contains hydrated silicates that are similar to some of minerals found in a typical carbonaceous chondrite matrix.

One way to compare micrometeorites directly with cometary dust is by using infrared measurements. Figure 8.13 shows the radiation which is transmitted when individual samples are illuminated by an appropriate infrared light source. The transmission spectra shown in Figure 8.13 are characteristic of the presence of certain chemical constituents: thus, pyroxene-rich particles can be distinguished from olivine-rich varieties (indeed, infrared measurements form the basis of a classification scheme for these materials). The infrared absorption features observed in micrometeorites can be compared with the infrared emission features of cometary dust. This is shown in Figure 8.14, where it can be seen that no particular class of micrometeorite matches the comet data. However, a closer

By 'chondritic composition' we mean here a chemical composition similar to CI carbonaceous chondrites or to the relevant part of such a chondrite – the CAIs and chondrules here.

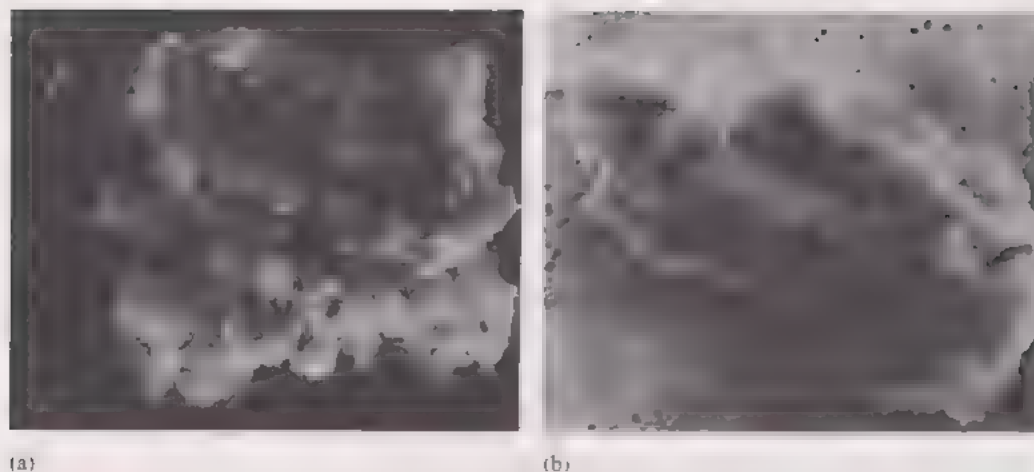


Figure 8.12 Scanning electron microscope pictures of micrometeorites. (a) An anhydrous particle, comprised of aggregates of smaller crystals. (b) A hydrated particle, showing a more compact and less porous morphology. It is composed predominantly of hydrated silicates. Both particles are about 10 μm across

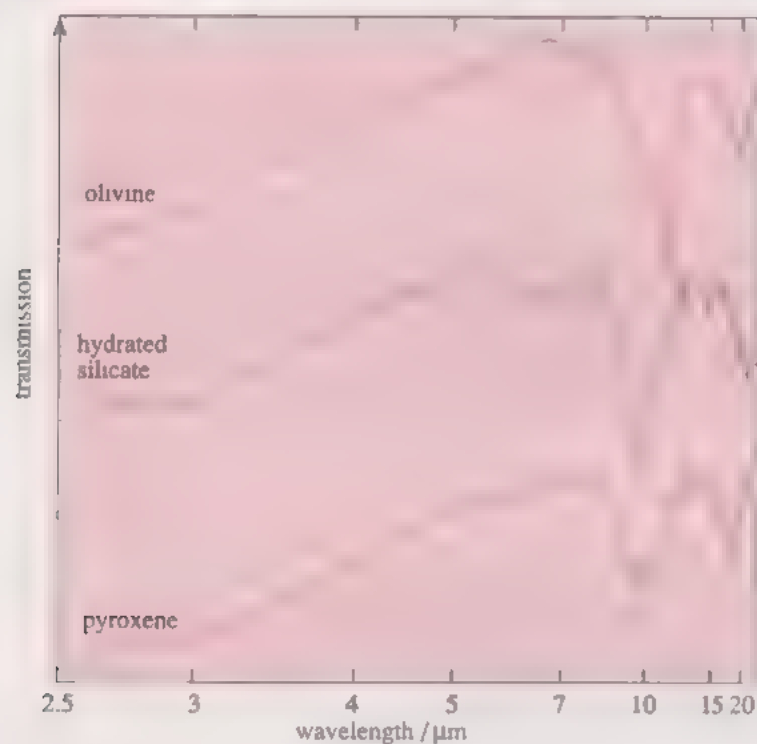


Figure 8.13 Infrared transmission spectra of three representative classes of micrometeorite, each dominated by a particular mineral type. (These are offset vertically to avoid overlap.) To the trained eye, such spectra can be used to classify particles of unknown type, notably in the 8–12 μm wavelength region (see Figure 8.14).

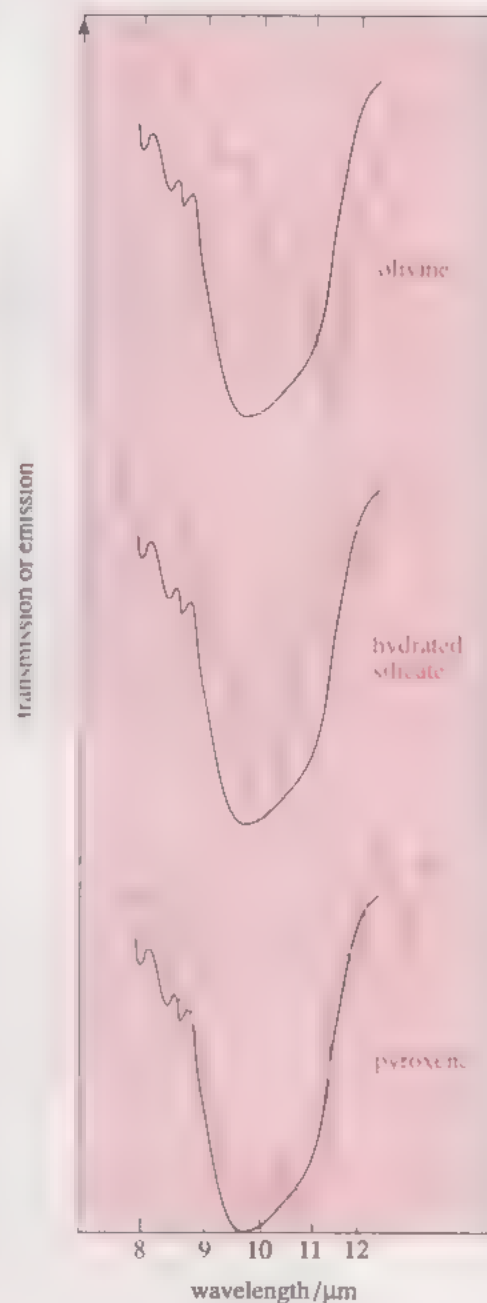


Figure 8.14 A comparison of infrared transmission spectra from three different classes of micrometeorite, each dominated by a particular mineral type (red curves) with the emission features obtained from telescopic observations of comet Kohoutek (black curves). The cometary data have been inverted to facilitate the comparison. (Each of the three pairs of curves is offset vertically for clarity.) Note that there is not a good match in any of the comparisons.

agreement can be obtained by using a composite of data. For instance, a reasonable explanation of the comet data is provided by a mixture of approximately equal amounts of pyroxene-rich particles and hydrated silicates. This is understandable since individual comets presumably contain a variety of different sorts of dust grains.

We have already stated that micrometeorites have an approximately chondritic composition. This shows that they are chemically similar to primitive meteorites. However, the high porosities and delicate textures observed in many micrometeorites are not represented in meteorites, implying that they are a very primitive class of extraterrestrial material. They do not match exactly any known type of meteorite.

ITQ 8.12 What do you think may have originally been in the pore spaces of micrometeorites, and how might the material in the pore spaces have been lost?

One of the features that distinguishes micrometeorites from meteorites is their high carbon contents (typically 5–15% by mass). When normalized to the content of silicon, the C/Si ratio (by number of atoms) in micrometeorites varies from 0.5 to 5. The upper end of this range coincides with data from Halley dust grains, which have C/Si of 4.4. In contrast, the C/Si ratio of the most carbon-rich primitive chondrite is only about 0.75. The carbon in the matrix of micrometeorites occurs in many forms including organic materials and poorly-crystalline graphite – these constituents appear to act as the embedding matrix which keeps the individual constituents of each micrometeorite together. Carbon also occurs in the form of iron–nickel carbide, silicon carbide and carbonates.

The origins of the carbonaceous materials are only partially understood. Carbonates, at least, are presumed to be a low-temperature product, possibly of secondary origin. On the other hand, the mass of amorphous and organic carbon, which occurs as rims on grains and filamentous components, appears to be formed by reactions involving gases, which presumably took place in the solar nebula. These materials could be related to the organic compounds in carbonaceous chondrites, but it is exceedingly difficult to assess this, even with the very latest analytical technology.

How might we try to understand the relationships of the carbon-bearing components in micrometeorites (to each other, or to materials observed in meteorites)?

■ Apart from studying the distribution of carbon in the particles, using microscopy, etc., it would be advantageous to determine the stable isotope compositions of the various components.

This is a difficult task, not least because the micrometeorites are very small. So far it has only been possible to determine the carbon isotope compositions of a few micrometeorites. In each case it was necessary to analyse the carbon from the entire sample, rather than from individual carbon-bearing components. The total range in $^{12}\text{C}/^{13}\text{C}$ is quite restricted (89 to 93), being similar to data from Earth, or whole meteorite samples. In contrast, $^{12}\text{C}/^{13}\text{C}$ ratios measured from Halley dust grains vary from 1–5 000. Thus, on the basis of the carbon isotope data, there would appear to be no obvious relationship between micrometeorites and Halley dust grains. However, this is an instance where the availability of only limited data may give a distorted view of the true picture. Furthermore, consider that micrometeorites have been brought into the terrestrial environment where they may have become contaminated by terrestrial carbon. Problems of this nature constitute an analytical challenge – the sort of thing that scientists relish!

Summary of Section 8.6 and SAQs

- 1 Giotto determined the dust flux emanating from comet Halley just after its perihelion passage. The upper end of the mass distribution of dust represents particles of about $10\text{ }\mu\text{m}$ in size.
- 2 Cometary (and asteroidal) dust is constantly added to and removed from the interplanetary medium. This dust produces features in the night sky known as zodiacal light and gegenschein.
- 3 Micrometeorites, of typical size $10\text{ }\mu\text{m}$, can be collected from the Earth's upper atmosphere. These particles can be studied in the laboratory thereby giving us an insight into the dusty parts of comets.
- 4 Micrometeorites are approximately chondritic in composition, but do not match exactly any of the known meteorite types.
- 5 High concentrations of carbon are found in micrometeorites, consistent with Halley dust grains. It is difficult to comprehend the nature and relationships of the carbon components in the particles because of analytical limitations.

SAQ 8.12 (Objective 8.12) In a hypothetical collection of dust from the upper atmosphere, how might you expect to distinguish micrometeorites of cometary origin from those that are asteroidal debris?

SAQ 8.13 (Objective 8.5) What geochemical measurements might help to prove that micrometeorites are not simply dust particles of terrestrial origin?

8.7 Conclusions

You have now come to the end of **Book 2, *The Planets***, in which we have described both the planets and the smaller bodies that make up our Solar System, and also outlined a theory of its origin. We have presented *some* of the evidence bearing on the origins of planetary bodies and for the structure and composition of their interiors and atmospheres, and have described a few of the phenomena that can be observed or inferred to occur. At the end of this last chapter we have been concerned with micrometeorites, which (molecules excluded) are some of the smallest objects studied by astronomers and planetary scientists. For **Book 3, *Galaxies***, we jump to the opposite end of the size spectrum, beginning first with our own galaxy and then moving on to consider other galaxies and the large-scale structure of the Universe.

Objectives for Chapter 8

After studying Chapter 8 (and any associated audio, video or TV material), you should be able to:

- 8.1 Give brief definitions of the terms, concepts and principles listed below.
- 8.2 Describe the broad categories of meteorites and appreciate that they represent fragments of various bodies from the Solar System.
- 8.3 Understand that meteorites constantly arrive on Earth and that certain environments, such as Antarctica, provide effective meteorite collection traps which can be exploited by dedicated retrieval parties.

- 8.4 Explain the relationship of the chemical composition of meteorites to that of the Sun.
- 8.5 Comprehend the relevance of oxygen isotope measurements to meteorite research and outline how these measurements are made.
- 8.6 Understand that certain primitive meteorites contain carbon in a variety of organic and elemental forms.
- 8.7 Appreciate that stable isotope measurements can be used to demonstrate the origin of certain components in meteorites.
- 8.8 Describe how the measurement of radioactive isotopes can constrain the formation ages of meteorites.
- 8.9 Understand that there are differences in the formation ages of meteorites and appreciate the relevance of this in relation to the evolution of the Solar System.
- 8.10 Understand why some meteorites are thought to come from Mars, and be able to describe their diagnostic features.
- 8.11 Comprehend certain aspects of the evolution of the Martian atmosphere.
- 8.12 Understand that micrometeorites mainly arise from comets and as such can be considered to have a somewhat different origin to that of meteorites.

List of scientific terms, concepts and principles used in Chapter 8

Term	Page	Term	Page	Term	Page
ablation	207	fireball	208	nakhlite	225
achondrite	206	fusion crust	207	ordinary chondrite	206
calcium–aluminium-rich inclusion	211	hydrodynamic escape	231	primitive	210
carbonaceous chondrite	206	hydrothermal alteration	211	rare-earth elements (REE)	221
chassignite	225	interplanetary dust	233	shergottite	225
chondrite	205	iron meteorite	205	spallation	218
chondrule	205	isotope exchange	215	stable isotopes	211
cosmic rays	218	matrix	215	stable isotope composition	211
cosmic ray exposure age	229	meteorite	204	SNC meteorites	226
cosmic sediments	211	meteor stream	209	stony meteorite	205
cosmic spherules	207	meteoroid	207	stony-iron meteorite	205
fall	204	micrometeorite	206	terrestrial fractionation line (TFL)	214
find	204	micrometeoroid	207	Widmanstätten pattern	224

ITQ answers and comments for Chapter 1

ITQ 1.1

The period can be found by using Equation 1.1, with $k = 1 \text{ yr}^2 \text{ AU}^{-3}$.

$$\tau^2 = (1 \text{ yr}^2 \text{ AU}^{-3}) \times (0.75 \text{ AU})^3$$

$$\tau = \sqrt{(0.75)^3 \text{ yr}^2}$$

$$= 0.65 \text{ yr, or 237 days}$$

[Comment: There is actually no planet at exactly 0.75 AU from the Sun. Venus is nearest; it is 0.722 AU distant, and has a period of 224.7 days, or 0.615 years.]

ITQ 1.2

The mass of Mars can be found by substituting appropriate values into the equation:

$$M + m = \frac{4\pi^2 a^3}{G\tau^2}$$

Ignoring the trivial mass of Deimos, this gives a value for M of

$$\frac{4\pi^2 (23\,460 \times 10^3 \text{ m})^3}{(6.67 \times 10^{-11} \text{ N m}^2 \text{ kg}^{-2}) \times (1.263 \times 86\,400 \text{ s})^2}$$

Making the substitution $1 \text{ N} = 1 \text{ kg m s}^{-2}$, this becomes

$$\begin{aligned} & \frac{4\pi^2 (23\,460 \times 10^3)^3 \text{ m}^3}{(6.67 \times 10^{-11} \text{ m}^3 \text{ kg}^{-1} \text{ s}^{-2}) \times (1.263 \times 86\,400)^2 \text{ s}^2} \\ &= 6.41 \times 10^{23} \text{ kg} \end{aligned}$$

Mars's density can also be easily obtained:

$$\rho = \frac{\text{mass}}{\text{volume}}$$

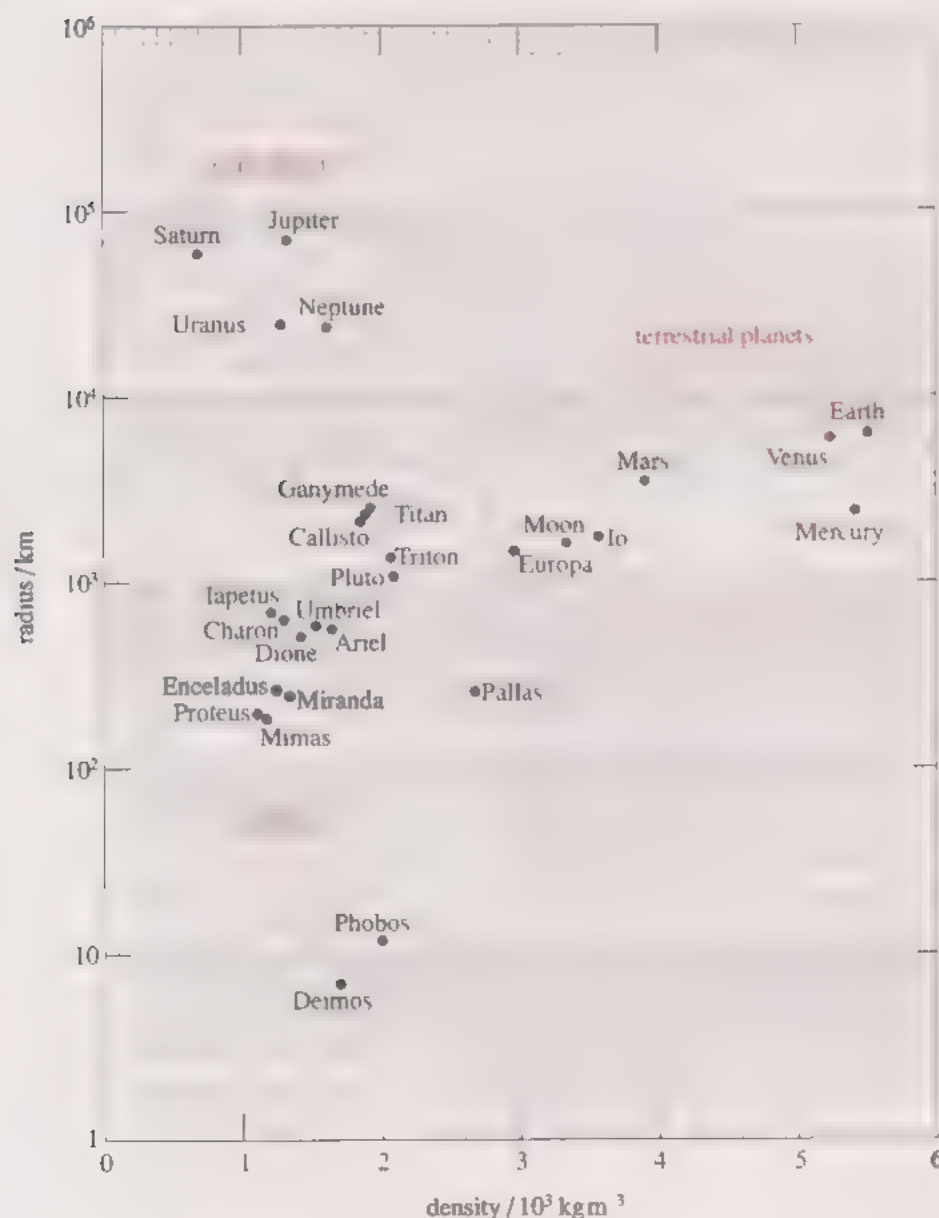


Figure 1.11 Completed version of Figure 1.4, for use with ITQ 1.3.

Thus

$$\begin{aligned}\rho &= \frac{6.41 \times 10^{23} \text{ kg}}{\frac{4}{3}\pi (3400 \times 10^3 \text{ m})^3} \\ &= 3.90 \times 10^3 \text{ kg m}^{-3}\end{aligned}$$

ITQ 1.3

Your plot should resemble that in Figure 1.11, where the main groups have been identified.

ITQ 1.4

Pluto, the smallest planet, is quite different in size, but not density, from its nearest neighbour, Neptune. Actually, in composition Pluto resembles the icy satellites – such as Neptune's largest satellite, Triton – much more closely than it does Neptune itself. It must be a mixture of rocky and icy materials.

ITQ 1.5

A 5:2 resonance implies that an asteroid would go five times round the Sun while Jupiter goes round twice. Thus, twice Jupiter's period is 2×11.86 years = 23.72 years, and one-fifth of this is $23.72/5 = 4.74$ years. The period of the asteroid would therefore be 4.74 years.

Using Kepler's third law, we can rapidly find how far from the Sun this would be (Section 1.2.1):

$$\tau^2 = ka^3, \text{ and so}$$

$$a^3 = \tau^2/k$$

The value of k was given in Subsection 1.2.1 as $1 \text{ yr}^2 \text{ AU}^{-3}$. Then, as in ITQ 1.1:

$$a^3 = \frac{(4.74 \text{ yr})^2}{1 \text{ yr}^2 \text{ AU}^{-3}} = 22.47 \text{ AU}^3$$

Thus $a = 2.82 \text{ AU}$. As Figure 1.6 shows, there is indeed a gap in this position.

ITQ answers and comments for Chapter 2

ITQ 2.1

The Solar System data given in Table 1.1a show that the inner limit of the Beta Pictoris dust cloud (20 AU) would lie within the orbits of Neptune (30.0 AU) and Pluto (39.4 AU), and fall close to the orbit of Uranus (19.1 AU). The outer limit of the Beta Pictoris disc (220 AU) lies at a distance greater than that of the orbit of Pluto, but far less than the distance from the Sun to the inner edge of the Oort cloud (20 000 AU) where most of the Sun's comets reside (Subsection 1.3.2).

ITQ 2.2

(a) As described by Kepler's third law (Equation 1.1), a particle in an inner orbit would have a shorter orbital period than one in an outer orbit.

(b) The slower outer part of the disc would tend to slow down the faster inner part, at the same time as the faster inner part would tend to speed up the slower outer part.

ITQ 2.3

The volume of a spherical planetesimal that has a diameter of 10 km, and thus a radius of 5 km, is:

$$\frac{4}{3} \times 3.14 \times (5 \times 10^3 \text{ m})^3 = 5.2 \times 10^{11} \text{ m}^3$$

The mass of an object is given by volume \times density, so the mass of such a planetesimal with the specified density is:

$$(5.2 \times 10^{11} \text{ m}^3) \times (3 \times 10^3 \text{ kg m}^{-3}) = 1.6 \times 10^{15} \text{ kg}$$

The number of such planetesimals necessary to make up the Earth is given by:

$$\begin{aligned}\frac{\text{mass of Earth}}{\text{mass of planetesimal}} &= \frac{6.0 \times 10^{24} \text{ kg}}{1.6 \times 10^{15} \text{ kg}} \\ &= 3.8 \times 10^9 \approx 4 \times 10^9\end{aligned}$$

ITQ 2.4

(a) Density decreases outwards from Jupiter, running from $3.57 \times 10^3 \text{ kg m}^{-3}$ for the innermost galilean satellite, Io, to $1.86 \times 10^3 \text{ kg m}^{-3}$ for the outermost galilean satellite, Callisto.

(b) The densities in Table 1.1b are mean densities for each satellite as a whole. The mean densities of Ganymede and Callisto are most simply explained if they are mixtures of rock and ice, with roughly equal proportions of each. The density of Europa can be explained by its being composed almost entirely of rock, but it would appear that we need to call on a denser component within Io (probably iron or iron and sulphur in its core). [Comment: Strictly speaking, we should take self-compression into account, but in bodies this small its effects are too slight to affect our argument.]

ITQ 2.5

The gravitational field of the planet would cause gravitational focusing, whereby the trajectories of incoming projectiles would be deviated towards the planet. As a result, the closer a satellite is to a planet, the more likely it is to be hit. A subsidiary factor is that the planet's gravity would accelerate incoming projectiles, so that those hitting an inner satellite would, on average, be travelling faster than those striking an outer satellite. The smaller the satellite, the less the energy required to break it up, so an incoming projectile carrying a given kinetic energy would be more able to fragment a small satellite.

ITQ answers and comments for Chapter 3

ITQ 3.1

(a) As in ITQ 2.3, the volume, V , of a sphere is given by $V = \frac{4}{3}\pi R^3$, and density is mass/volume. Substituting the given values, we get

$$\begin{aligned}\text{bulk density of Earth} &= \frac{\text{mass of Earth}}{\text{volume of Earth}} \\ &= \frac{5.98 \times 10^{24} \text{ kg}}{\frac{4}{3}\pi(6.37 \times 10^6 \text{ m})^3} = \frac{5.98 \times 10^{24} \text{ kg}}{1.08 \times 10^{21} \text{ m}^3} \\ &= 5.53 \times 10^3 \text{ kg m}^{-3}\end{aligned}$$

(b) Rocks sampled in the continents and on the ocean floor have only about half the density of the bulk Earth. This means that the Earth must become denser deeper down, and so must contain a zone that is denser than the average value in order to compensate for the less dense rock near the surface.

ITQ 3.2

S-wave speed rises suddenly from zero (i.e. non-transmission) to about 3.6 km s^{-1} at a depth of about 5 150 km. This depth must mark the limit of the inner core; note that there is a rise in P-wave speed at this depth also. Since the Earth's radius is 6 370 km, then the radius of the inner core must be $\approx (6\,370 - 5\,150) \text{ km} \approx 1\,220 \text{ km}$.

ITQ 3.3

The model there is that terrestrial planets grew by collisions between planetary embryos, over a period of 10^7 to 10^8 years. These embryos would have had similar compositions, because they came from a short range of distances from the Sun. It is conceivable that the planetary embryos had accreted heterogeneously, but there would be no way for the planets themselves, if formed by collisions between embryos, to be built up one compositional layer at a time. Thus, if the model is correct, the terrestrial planets were formed by homogeneous accretion. [Comment: There is a school of thought that the Earth's oceans were provided by late accretion of ice-rich cometary debris, which requires an episode of heterogeneous accretion right at the end.]

ITQ 3.4

The nickel-iron would sink towards the centre of the planet. [Comment: In fact this would happen even if the temperature were adequate to melt the nickel-iron as well, because this would form liquid globules that would not mix with the silicate melt, but would sink gravitationally in the same manner as solid chunks. See for yourself by gently stirring a small glass containing two liquids of different density that do not mix easily, and watch what happens. We recommend a dense liquid such as grenadine syrup, and a less dense liquid such as tequila. This is called a tequila sunrise, and is very tasty!]

ITQ 3.5

(a) It is to the near-surface region that the energy is added by collision. This is conducted only slowly towards the interior.

(b) As the planet grows more massive its gravitational attraction becomes correspondingly greater. Incoming projectiles are thus accelerated to greater speeds, and so they carry more kinetic energy upon impact, which is then converted into heat.

ITQ 3.6

The number of seconds in a year is 3.2×10^7 . So the rate of heat generation at $3 \times 10^{19} \text{ J yr}^{-1}$ is given by

$$\frac{3 \times 10^{19} \text{ J}}{3.2 \times 10^7 \text{ s}} = 9 \times 10^{11} \text{ J s}^{-1} = 9 \times 10^{11} \text{ W}$$

The rate of tidal heat production per kilogram is given by

$$\begin{aligned}\frac{\text{total rate of tidal heat production}}{\text{mass of the Earth}} &= \frac{9 \times 10^{11} \text{ W}}{6 \times 10^{24} \text{ kg}} \\ &= 1.5 \times 10^{-13} \text{ W kg}^{-1} \approx 1.5 \times 10^{-10} \text{ mW kg}^{-1}\end{aligned}$$

The present rate of radiogenic heat production within the Earth, obtained by summing the four values quoted in Table 3.2, is about $5 \times 10^{-9} \text{ mW kg}^{-1}$, which is about 30 times greater than the rate of tidal heating calculated here.

ITQ answers and comments for Chapter 4

ITQ 4.1

The proportion of elliptical craters is negligible. Given the wide range of possible impact angles, this shows that the process of cratering favours production of circular craters over elliptical ones.

ITQ 4.2

Because the crater is so tiny, the meteorite was clearly a very small one. Its small size means that its speed must have been severely limited by atmospheric drag. Thus, it more nearly resembled a pebble thrown into mud than an explosive hyper-velocity impact. The asymmetrical distribution of the ejecta suggests that the meteorite was travelling from right to left. Thus, it probably struck at an oblique angle, greater than 45° from the perpendicular.

ITQ 4.3

On Earth, thick masses of ice deform under their own weight, so that they *flow* in the solid state, forming glaciers. If ice could flow on planetary satellites, the rugged rims of craters would gradually fade away.

ITQ 4.4

Figure 4.14b (an image of the heavily cratered terrain on Mercury) shows a nearly saturated surface – it would be impossible to fit a large new crater on it without overprinting on older ones. In fact, some partial overprinting has already occurred. [*Comment:* The large crater at the top left (Rabelais) is 120 km across. In detail, the surface is not fully saturated because there are areas of ‘inter-crater plains’ where it would be possible to insert craters 15–20 km in diameter without obliterating others.] Figure 4.14a shows an area (of ‘smooth plains’ on Mercury) superficially similar to the lunar maria. It is clearly not saturated, and must be younger than that shown in Figure 4.14b.

ITQ 4.5

There are two parts to the curve (Figure 4.22). For crater diameters greater than about 24 km, there is a roughly linear relationship (on the logarithmic graph) between crater diameter and frequency. One would expect there to be far more small craters than large, but the curve does not show this. The curve undulates a bit, but instead of rising towards the smaller end, there is a drop-off in the frequency of craters smaller than about 24 km.

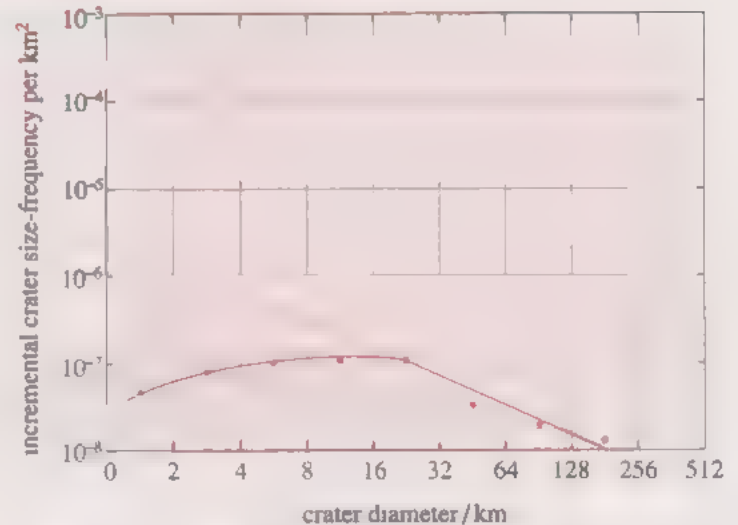


Figure 4.22 Answer for ITQ 4.5. Size–frequency distribution curve for terrestrial craters.

ITQ 4.6

Extrapolating the rate of 5.6×10^{-5} craters km^{-2} per billion years for 3.2 billion years yields 1.8×10^{-4} craters km^{-2} . Multiplying by the area of the mare surface gives:

$$1.8 \times 10^{-4} \text{ km}^{-2} \times 200\,000 \text{ km}^2 \text{ craters, or } 36 \text{ craters}$$

ITQ 4.7

The production rate on the lunar maria for 32–64 km craters over the last 3.2 billion years ($1.5 \times 10^{-6} \text{ km}^{-2}$) averages out at around $4.7 \times 10^{-7} \text{ km}^{-2}$ per billion years, or $4.7 \times 10^{-7} \text{ km}^{-2} \times 10^{-9}$ per year; equal to $4.7 \times 10^{-16} \text{ km}^{-2} \text{ yr}^{-1}$. On the Canadian Shield, there are about 2×10^{-7} craters km^{-2} . Thus, dividing the observed crater density by the rate of cratering, the age of the Canadian Shield ought to be about:

$$\frac{2 \times 10^{-7} \text{ km}^{-2}}{4.7 \times 10^{-16} \text{ km}^{-2} \text{ yr}^{-1}} \approx 4.3 \times 10^8 \text{ yr}$$

This is about 430 million years. [*Comment:* Many of the rocks making up the Canadian Shield are far older, but it is the age of the *surface* that concerns us. Even this is a little oversimplified, because erosion has removed a significant veneer from the original surface.]

ITQ 4.8

Your completed table should resemble Table 4.4. Mars has clearly had a rich geological history, the activity persisting longer and being more diverse than the Moon's.

Table 4.4 Crater densities of some Martian geological terrains

Geological province	Crater density relative to that of average lunar mare	Estimated age / 10^9 yr
Tharsis volcanic plains	0.1	0.4
Elysium volcanic plains	0.7	3.0
Chryse Planitia	1.1	3.4
Tyrrenum Patera	1.4	3.5
Hellas	1.8	3.6
Syrts Major	2.0	3.7
Heavily cratered plains	13	4.1

ITQ 4.9

Your curve should resemble Figure 4.23. The curve is similar to that for the lunar highlands for the largest craters, since these would be least obliterated by erosion, which eliminates small craters preferentially. At the smallest size end, the curve should resemble that for the lunar maria. Between the two, the curve should flatten off, representing the population of 'lost' craters. [Comment: This is why we included data only for large craters for the heavily cratered plains in Table 4.3.]

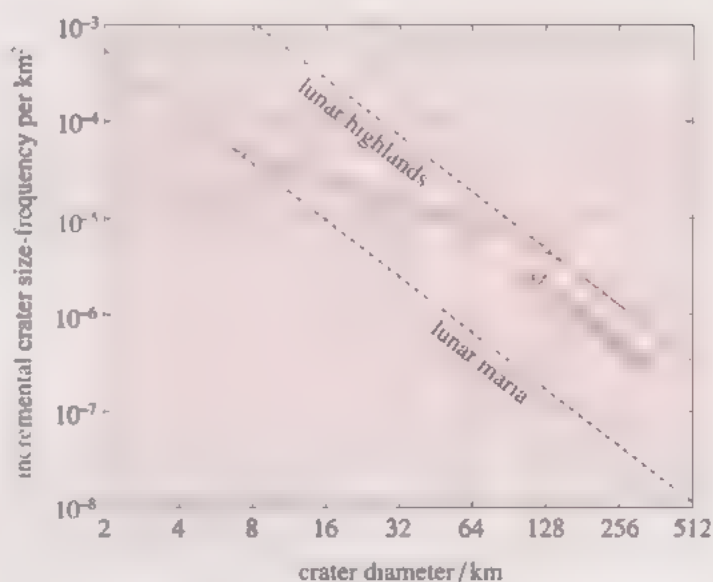


Figure 4.23 Answer for ITQ 4.9. Size-frequency distribution curve for Mars's highland craters.

ITQ answers and comments for Chapter 5

ITQ 5.1

The lunar mare lava and komatiite lava have the most MgO, and are much closer to 'primitive' basalts and fresh mantle than the Martian lava, which contains less MgO and more silicon, sodium and potassium, and is more MORB-like. As outlined in Subsection 5.2.1, partial melting leads to increases in the concentration of these elements in the liquids (lavas) that result. Thus, the composition of the Martian lava suggests that fractionation processes similar to those producing MORBs has taken place.

ITQ 5.2

Yes – there is good evidence for orbital resonance. Enceladus's orbital period is almost exactly half that of Dione, and one and a half times that of Mimas. [Comment: These orbital resonances were the subject of SAQ 3.4. Furthermore, the orbital period of Enceladus is also close to

double that of two very tiny satellites, Janus and Epimetheus (both less than about 100 km in radius). At some point in the past, Enceladus's orbit may have been in exact 4:2:1 resonance with the orbits of Dione and Janus and Epimetheus. In this situation, sufficient heat may have been generated by tidal dissipation to drive volcanism, causing the resurfacing evident in the image.]

ITQ 5.3

(a) g , the acceleration due to gravity (at the surface), appears as a variable in Equation 5.1. Because g appears in the denominator, lower values of g mean higher values of h . Thus, the thickest flows will be found on asteroids, which have the lowest values of g .

(b) g for the Moon is 1.6 m s^{-2} ; for Mercury it is 3.7 m s^{-2} . Because the other variables remain the same, flow

thickness is inversely proportional to g . Thus, the thickness of the flow on Mercury would be:

$$\frac{1.6}{3.7} \times 10 \text{ m} = 4.3 \text{ m}$$

ITQ 5.4

This is basically a question about terminal speed. In this case, the terminal fall speed must be just less than the upwards speed of the escaping gas. Use Equation 5.3.

$$\text{terminal speed} = C_d \sqrt{\frac{(dg\rho)}{\beta}}$$

Thus, inserting values:

$$100 \text{ m s}^{-1} = 1.054 \times \sqrt{\frac{(d \times 9.8 \text{ m s}^{-2} \times 1.5 \times 10^3 \text{ kg m}^{-3})}{1.2 \text{ kg m}^{-3}}}$$

Squaring both sides:

$$(100 \text{ m s}^{-1})^2 = (1.054)^2 \times \frac{(d \times 9.8 \text{ m s}^{-2} \times 1.5 \times 10^3 \text{ kg m}^{-3})}{1.2 \text{ kg m}^{-3}}$$

Multiplying out brackets on each side:

$$1.0 \times 10^4 \text{ m}^2 \text{ s}^{-2} = (1.36 \times 10^4 \text{ m s}^{-2}) \times d$$

Thus

$$d = \frac{1.0 \times 10^4 \text{ m}^2 \text{ s}^{-2}}{1.36 \times 10^4 \text{ m s}^{-2}} = 0.74 \text{ m}$$

[*Comment:* This sounds like a large lump of pumice, but it is consistent with what is observed in major eruptions, near the vent.]

ITQ 5.5

Convection is impossible on a planet without an atmosphere, because there is nothing to provide buoyant support to the column, and no gas to be entrained and heated. Erupted gases therefore expand into the vacuum of space, driving ejected fragments along ballistic trajectories.

ITQ 5.6

Mars is a small silicate planet, lacking major satellites. Its only *present-day* source of heat therefore must be from decay of long-lived radioactive isotopes, plus whatever remains of its primordial heat. Because it is small – its mass is only about one-tenth that of the Earth – its active volcanic lifetime is likely to have been short, but not as short as that of the Moon. Like the Earth and Moon, one would expect to find abundant evidence of lava eruptions, particularly in the *early* history of the planet. Since Mars is rich in volatiles (Chapter 4), there should be evidence of pyroclastic eruptions. Because Mars has a thin atmosphere, convecting eruption columns would be possible.

ITQ 5.7

Although Io and Earth have similar heat productions, Io has only one-twelfth of the surface area of the Earth (the ratio of the squares of the radii, from Tables 1.1a and b). Thus, Io's heat production drives a much greater heat flow per unit area of surface, and more rapid melting processes.

ITQ answers and comments for Chapter 6

ITQ 6.1

From Equation 6.1 ($m_c = P/g$) and using the values given, for Venus:

$$\begin{aligned} P &= 92 \text{ bar} \\ &= 9.2 \times 10^6 \text{ Pa} \\ &= 9.2 \times 10^6 \text{ kg m}^{-1} \text{ s}^{-2} \end{aligned}$$

$$\begin{aligned} m_c &= \frac{9.2 \times 10^6 \text{ kg m}^{-1} \text{ s}^{-2}}{8.90 \text{ m s}^{-2}} \\ &= 1.03 \times 10^6 \text{ kg m}^{-2} \end{aligned}$$

and for Mars,

$$\begin{aligned} P &= 6.3 \times 10^{-3} \text{ bar} \\ &= 6.3 \times 10^2 \text{ kg m}^{-1} \text{ s}^{-2} \end{aligned}$$

$$\begin{aligned} m_c &= \frac{6.3 \times 10^2 \text{ kg m}^{-1} \text{ s}^{-2}}{3.72 \text{ m s}^{-2}} \\ &= 1.69 \times 10^2 \text{ kg m}^{-2} \end{aligned}$$

The column mass for Venus is about 100 times that for Earth, and the column mass for Mars is about 1/60 of that for Earth.

ITQ 6.2

The geothermal flux given in Subsection 3.5.2 is 0.08 W m^{-2} averaged over the Earth's surface. The total power from this source is $4\pi R^2 \times 0.08 \text{ W m}^{-2}$. The total solar flux absorbed is πR^2 times the solar flux density, multiplied by $(1 - a)$, which is the fraction absorbed.

The ratio of these is:

$$\begin{aligned} &\frac{\pi R^2 \times (1.38 \times 10^3 \text{ W m}^{-2}) \times (1 - 0.23)}{4\pi R^2 \times 0.08 \text{ W m}^{-2}} \\ &= \frac{1.38 \times 10^3 \times 0.77 \text{ W m}^{-2}}{4 \times 0.08 \text{ W m}^{-2}} \\ &= 3.3 \times 10^3 \end{aligned}$$

[*Comment:* Geothermal flux thus accounts for only 0.03% of the power input to the Earth's surface and atmosphere. In considering the atmosphere, this input is not significant. Similar conclusions are reached for Venus and Mars. So it is valid to compare effective temperatures on these planets as a measure of the power re-radiated by the atmospheres after absorption of solar radiation.]

ITQ 6.3

From Wein's law, at a temperature of 288 K

$$\begin{aligned}\lambda_{\text{peak}} &= \frac{2.90 \times 10^{-3} \text{ m K}}{288 \text{ K}} \\ &= 1.01 \times 10^{-5} \text{ m}\end{aligned}$$

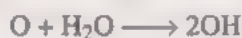
Electromagnetic radiation of this wavelength is in the infrared region (Book 1, Subsection 1.4.2). So the Earth's radiation is most intense in the infrared region.

ITQ 6.4

Because (i) r decreases by the factor $\cos 30^\circ$ and (ii) mvr , the magnitude of angular momentum, is constant, then v increases by $1/(\cos 30^\circ)$. So, at 30° N , $v = 537 \text{ m s}^{-1}$.

ITQ 6.5

The balanced equation is



[*Comment:* The reactants contain two atoms of oxygen and two atoms of hydrogen, and so must the products. As the product is hydroxyl (OH), two molecules of OH must be produced from the reaction of each O atom and H₂O molecule. This is depicted by writing a 2 in front of OH in the equation.]

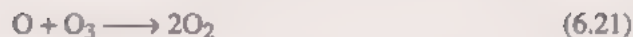
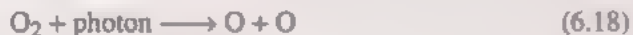
ITQ 6.6

(i) X is NO. [*Comment:* The effect of O (Figure 6.27) on N₂O is to oxidize it, via $\text{N}_2\text{O} + \text{O} \longrightarrow 2\text{NO}$. The molecule NO is reactive and is quickly converted into NO₂.]

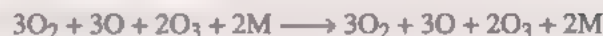
(ii) Y is OH. [*Comment:* OH is the common oxidizing substance in the atmosphere. The reaction is $\text{NO}_2 + \text{OH} \longrightarrow \text{HNO}_3$.]

ITQ 6.7

The net effect is given by the summation of the equations, although it is necessary to count Equation 6.19 twice in order to take account of the fact that one molecule of ozone O₃ is consumed in each of the reactions 6.20 and 6.21 (this also takes care of the two oxygen atoms produced in Equation 6.18):



The summation is thus



and there is no net effect. [*Comment:* The Chapman scheme (along with other reactions involving O₃), maintains a balance between the rate of formation of O₃ and its rate of destruction.]

The presence of M in Equation 6.19 is necessary because a bond is formed in this reaction and, to dispose of the excess energy, some energy must be transferred to the third body M.

ITQ 6.8

The atmosphere of Venus at cloud level and above is very dry (see Table 6.6) owing to the dehydrating action of H₂SO₄ droplets and the chemistry involved in the formation of H₂SO₄. In the near absence of H₂O, the concentration of OH radicals will be very low.

ITQ 6.9

The volume of the mantle is the volume of the Earth minus the volume of the sphere (i.e. the core) lying below the mantle:

$$\begin{aligned}V_{\text{mantle}} &= \frac{4}{3}\pi\{(6.37)^3 - (3.37)^3\} \times 10^{18} \text{ m}^3 \\ &= 9.22 \times 10^{20} \text{ m}^3\end{aligned}$$

The mass of the mantle is given by its volume times its density:

$$\begin{aligned}M_{\text{mantle}} &= V_{\text{mantle}} \times \rho_{\text{mantle}} \\ &= (9.22 \times 10^{20} \text{ m}^3) \times (4.0 \times 10^3 \text{ kg m}^{-3}) \\ &= 3.69 \times 10^{24} \text{ kg}\end{aligned}$$

0.1% of this mass is $3.69 \times 10^{21} \text{ kg}$. [*Comment:* This is about 6×10^{-4} of the mass of the Earth, which is about the same as the *total* estimated volatile inventory of the Earth! It indicates the difficulty of estimating volatile inventories for planets.]

ITQ answers and comments for Chapter 7

ITQ 7.1

For Enceladus, $a = 238\,000\text{ km}$, and $\tau = 1.370 \times 24 \times 3\,600\text{ s}$.

Using the equation from Chapter 1,

$$M = \frac{4\pi^2 a^3}{G\tau^2}$$

$$\begin{aligned} M_{\text{Saturn}} &= \frac{4\pi^2 (238\,000 \times 10^3\text{ m})^3}{(6.67 \times 10^{-11}\text{ N m}^2\text{ kg}^{-2})(1.370 \times 24 \times 3\,600\text{ s})^2} \\ &= \frac{4\pi^2 \times (1.348 \times 10^{25}\text{ m}^3)}{(6.67 \times 10^{-11}\text{ N m}^2\text{ kg}^{-2})(1.401 \times 10^{10}\text{ s}^2)} \\ &= 5.69 \times 10^{26} (\text{m}^3)(\text{N}^{-1}\text{ m}^{-2}\text{ kg}^2)(\text{s}^{-2}) \end{aligned}$$

Now

$$1\text{ N} = 1\text{ kg m s}^{-2}, \text{ so } 1\text{ N}^{-1} = 1\text{ kg}^{-1}\text{ m}^{-1}\text{ s}^2.$$

Therefore

$$\begin{aligned} M_{\text{Saturn}} &= 5.69 \times 10^{26} (\text{m}^3)(\text{kg}^{-1}\text{ m}^{-1}\text{ s}^2\text{ m}^{-2}\text{ kg}^2)(\text{s}^{-2}) \\ &= 5.69 \times 10^{26}\text{ kg} \end{aligned}$$

ITQ 7.2

Density is mass per unit volume. The approximate volume of Saturn was calculated as $9.0 \times 10^{23}\text{ m}^3$. Hence the density of Saturn is approximately

$$\frac{5.69 \times 10^{26}\text{ kg}}{9.0 \times 10^{23}\text{ m}^3} = 6.3 \times 10^2\text{ kg m}^{-3}$$

[*Comment:* This is a very low density. It is lower than the density of water at the surface of the Earth ($1\,000\text{ kg m}^{-3}$).]

ITQ 7.3

The heat produced by radiogenic decay on Earth is $4.8 \times 10^{-9}\text{ mW kg}^{-1} = 4.8 \times 10^{-9} \times 10^{-3}\text{ W kg}^{-1}$. Eight Earth masses is $5.98 \times 10^{24} \times 8\text{ kg}$, so a core of eight Earth masses on Jupiter would produce

$$\begin{aligned} &(4.8 \times 10^{-9} \times 10^{-3}\text{ W kg}^{-1}) \times (5.98 \times 10^{24} \times 8\text{ kg}) \\ &= 2.3 \times 10^{14}\text{ W} \end{aligned}$$

To obtain the flux we have to divide by the surface area of Jupiter. As an approximation, we take the emitting layer as the 1 bar layer and the planet as spherical. [*Comment:* Different wavelengths are emitted from different depths in the atmosphere, but infrared radiation all comes from layers close to 1 bar and the distance of these layers from the 1 bar level will be negligible compared with the radius of Jupiter.]

The approximate surface area is thus $4\pi R_J^2 = 4\pi \times (7.14 \times 10^7\text{ m})^2$. The flux is thus

$$\frac{2.3 \times 10^{14}\text{ W}}{4\pi \times 5.10 \times 10^{15}\text{ m}^2} = 3.6 \times 10^{-3}\text{ W m}^{-2}$$

This is far too small to account for the observed flux (7 W m^{-2}).

ITQ 7.4

Figure 7.23 shows the vertical profile for Titan with the labels troposphere (this is the region near the surface where the temperature falls with altitude) and thermosphere (where the temperature increases with altitude). The curve is similar in shape to that for Mars (Figure 6.6).

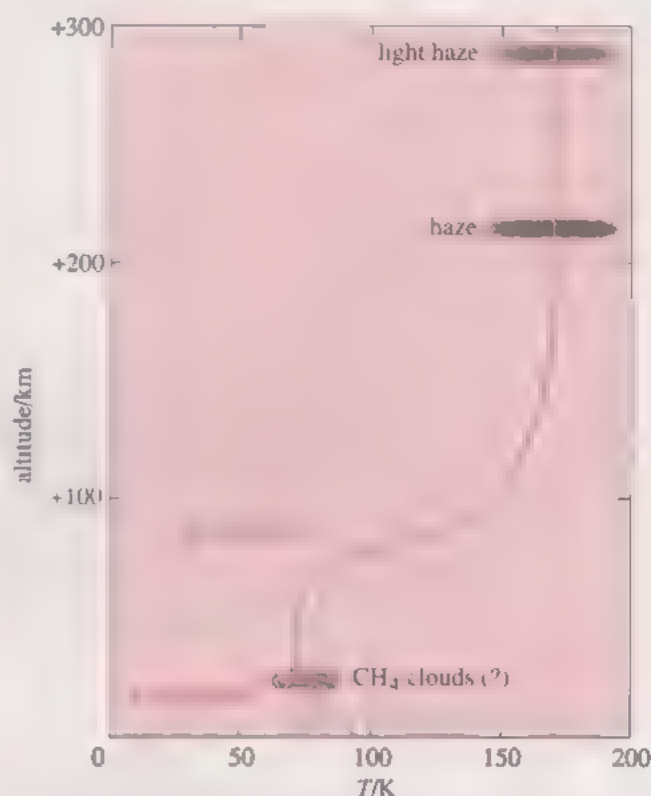


Figure 7.23 Vertical profile of Titan's atmosphere with the regions labelled

ITQ 7.5

A still atmosphere means that there is no wind and so it is rotating with the centre of the planet with a period of 0.410 days. Thus at the equator it tracks the circumference of the planet in 0.410 days. The circumference is given by $2\pi R$, where R is the equatorial radius. Hence the 'still' atmosphere travels $2\pi \times 71\,400\text{ km}$ in 0.410 days, that is $2\pi \times 71\,400\text{ km}$ in $0.410 \times 24\text{ h}$. This is $45\,600\text{ km h}^{-1}$, or 12.7 km s^{-1} , or $12\,700\text{ m s}^{-1}$. The direction is west to east – that of the planet's rotation.

ITQ answers and comments for Chapter 8

ITQ 8.1

The total mass of material added is equal to the yearly flux multiplied by the length of time over which material has been added. Thus

$$\text{total mass} = (10^8 \text{ kg yr}^{-1}) \times (4 \times 10^9 \text{ yr}) = 4 \times 10^{17} \text{ kg}$$

The volume of material added is given by

$$\frac{\text{mass}}{\text{density}} = \frac{4 \times 10^{17} \text{ kg}}{1.5 \times 10^3 \text{ kg m}^{-3}} = 2.7 \times 10^{14} \text{ m}^3$$

The surface area of a sphere is $4\pi R^2$. Thus, the volume of the layer is $4\pi R_{\oplus}^2 \times h$, where R_{\oplus} is radius of the Earth and h is the depth of the layer. So

$$2.7 \times 10^{14} \text{ m}^3 = 4\pi(6.38 \times 10^6)^2 \text{ m}^2 \times h$$

Therefore

$$h = \frac{2.7 \times 10^{14} \text{ m}^3}{5.1 \times 10^{14} \text{ m}^2} = 0.53 \text{ m}$$

ITQ 8.2

The alteration has taken place in a closed system. Although fluids have been formed and these have reacted with the original minerals to form new phases, no chemicals have been carried away, and none has been brought in from elsewhere. [Comment: Presumably, there was not a great deal of fluid flow through the parent body, otherwise the fluids would have taken up certain elements into solution, or suspension, and removed them to another part of the parent body, thereby changing the chemical composition of the small samples represented by individual meteorites.]

ITQ 8.3

The $\delta^{18}\text{O}$ of ocean water is $\left(\frac{1/499}{1/499} - 1\right) \times 1000 = (1 - 1) \times 1000 = 0\text{‰}$. Similarly $\delta^{17}\text{O}$ can be shown to be 0‰ . In other words, the reference point always has a δ value of 0‰ . Relative enrichments of the isotopes ^{17}O and ^{18}O result in δ values that have a positive sign. Pure ^{16}O would have a $\delta^{18}\text{O}$ of $\left(\frac{0/1}{1/499} - 1\right) \times 1000 = (0 - 1) \times 1000 = -1000\text{‰}$, and the $\delta^{17}\text{O}$ value of pure ^{16}O works out at the same value.

ITQ 8.4

Remember that pure ^{16}O has $\delta^{17}\text{O}$ and $\delta^{18}\text{O}$ of -1000‰ . Starting at G and adding ^{16}O to the system would produce a line with a slope of about 1 on Figure 8.4 leading away from G towards $-1000, -1000$. (If you have trouble with understanding this, draw an expanded graph with axes that go from -1000 to 0 for each δ value. Pure ^{16}O occurs at $-1000, -1000$, while pure G is approximately at $0, 0$. Mixing between these points produces a line of slope close to 1.) [Comment: When the oxygen isotope data from CAIs in Allende were first obtained, the addition of pure ^{16}O to

materials with an isotope composition somewhere in the vicinity of G was considered a reasonable explanation of the data (the ^{16}O was assumed to have been injected into the solar nebula at a late stage from a supernova). However, for this theory to have gained acceptance it was considered that further oxygen isotope measurements would have eventually uncovered more extreme values than point D, and ideally all the way down to $-1000, -1000$. Thus far, innumerable measurements have failed to exceed the limit set by D, which seems to suggest that this was the isotope composition of a large reservoir of dust.]

ITQ 8.5

The simple answer is that planets and asteroids that formed at different times will have different oxygen isotope compositions. [Comment: Note that some of the differences may be related to the place rather than to the timing of formation.]

ITQ 8.6

The survival of interstellar organic molecules shows us that there must have been places within the solar nebula that were never heated significantly above 400 K . [Comment: Although at the centre of the nebula, in the region of the Sun, temperatures were high, the outer parts, where comets formed, remained relatively cold. Carbonaceous chondrites sample materials from both extremes, demonstrating considerable mixing in the nebula.]

ITQ 8.7

It is most probable that the pre-solar relicts have been almost entirely assimilated into a Solar System host. For instance, pre-solar ^{16}O -rich oxide minerals may have reacted to an unknown extent with gases in the solar nebula thereby changing their original isotope composition. If these materials were isolated in their purest form, it would not be possible to state that they were entirely pre-solar. Similarly, deuterium-rich organic compounds may have become irreversibly associated with organics of Solar System origin.

ITQ 8.8

By measuring the content of ^{26}Mg . Because some minerals contain Al and Mg it is difficult in these instances to measure any excess ^{26}Mg produced by decay of ^{26}Al . It is desirable to search for the effects of ^{26}Al decay in minerals which originally would have contained very little, or no, magnesium.

ITQ 8.9

Electromagnetic induction and heating through impact events (if not so violent that the asteroid is permanently disrupted) are possibilities. However, the decay of short-lived radioactive isotopes, such as $^{26}_{13}\text{Al}$, is the most likely process to have contributed heat (see Subsection 8.4.2).

ITQ 8.10

Mars has an atmosphere and shows evidence of erosion by flowing water (Figure 4.19). In addition, winds transport surface materials and act as a weathering agent. The surface of Mars can be considered to be relatively active compared with that of the Moon. Basalts collected from the Moon are in much the same condition as when they were erupted. On the other hand, it may be anticipated that meteorites from Mars would record something of the active surface of the planet.

ITQ 8.11

We could measure the $^{12}\text{C}/^{13}\text{C}$ ratios of CO_2 from the glass fraction of EET A79001 and also from unshocked SNC meteorites. If there is no substantial difference in the two we could conclude that CO_2 had not been lost from the Martian atmosphere. [Comment: In fact, when we do this we find that the $^{12}\text{C}/^{13}\text{C}$ ratio of atmospheric CO_2 is 86, whereas that for the unshocked sources is 91. Errors on

these numbers are insignificant compared with the difference between them. It appears from the data acquired from SNC meteorites that the Martian atmosphere has become enriched in ^{13}C over the course of time, most probably by the preferential loss of the lighter ^{12}C isotope to outer space.]

ITQ 8.12

Comets are mixtures of ice and dust. Thus, when originally in the comet, the pore spaces in the micrometeorite would have been filled with ices (mainly water-ice). These would have been removed, either during ejection of the particle from the comet (accompanied by vaporization of volatiles), or by frictional heating during its fall through the atmosphere.

SAQ answers and comments for Chapter 1

SAQ 1.1

Chiron's period is found using Kepler's third law as in ITQs 1.1 and 1.5:

$$\tau^2 = ka^3$$

Therefore

$$\tau^2 = (1 \text{ yr}^2 \text{ AU}^{-3}) \times (13.7 \text{ AU})^3 = 2571 \text{ yr}^2$$

and thus

$$\tau = 50.7 \text{ yr}$$

Because Chiron is so small and distant from the Sun, it is most likely to be a low-density, icy body, similar in composition to the small satellites of the giant planets, or a comet. [Comment: In 1988, telescopic observations revealed that Chiron has a faint 'coma' of escaping gases, confirming its cometary nature.]

SAQ 1.2

(a) Comet Halley is extremely small, about the same size as Phobos and Deimos, but much lower in density. It therefore plots towards the bottom of the diagram to the left of Phobos and Deimos. These two satellites are its nearest parallels, but it probably does not resemble them closely. Some Earth-crossing asteroids may be 'dead' comets, which have lost their ice, but there is no reason to suppose that Phobos and Deimos, captured from the main asteroid belt, were ever icy.

(b) The data in Table 1.1c suggest that Vesta should have a density of very roughly $3 \times 10^3 \text{ kg m}^{-3}$. It thus plots well

above Phobos and Deimos in terms of size, and its density approaches that of the 'terrestrial' planets. [Comment: If there were enough asteroid data, they would probably define a field extending from Phobos and Deimos towards Earth.]

Using Equation 1.3, and the values of mass and radius in Table 1.1c, the escape speed from Vesta is given by

$$\begin{aligned} v_{\text{esc}} &= \left[\frac{2 \times (6.67 \times 10^{-11} \text{ N m}^2 \text{ kg}^{-2}) \times (2 \times 10^{20} \text{ kg})}{2.5 \times 10^5 \text{ m}} \right]^{1/2} \\ &= 330 \text{ m s}^{-1} \\ &= 0.33 \text{ km s}^{-1} \end{aligned}$$

A lower limit for Vesta's surface temperature can be obtained from Table 1.1a, noting from Table 1.1c that Vesta orbits the Sun between Mars and Jupiter. Vesta should certainly be much warmer than Pluto at its warmest – i.e. much warmer than -165°C , or 108 K. From Figure 1.10 it is then clear that Vesta should have little or no atmosphere.

SAQ 1.3

An Earth-like planet would orbit an alien sun at the same distance and with the same orbital period *only if that sun had the same mass as ours*. The orbital periods of the planets are dictated by the gravitational field set up by the Sun (Subsection 1.2.2).

SAQ answers and comments for Chapter 2

SAQ 2.1

The disc shape is the result of conservation of angular momentum, which inhibits collapse in the equatorial plane, but has no effect on matter collapsing from above the poles. In the hot, dense inner part of the disc, collapse is also inhibited by gas pressure. [Comment: Collapse may also be inhibited by magnetic forces, which we do not consider in this book.]

SAQ 2.2

Viscous drag within the solar nebula while the protoSun was forming will have led to outward transfer of angular momentum. The material falling into the protoSun would tend to be that whose speed of rotation had been slowed down by viscous drag. The T Tauri solar wind, which occurred later (we shall consider *when* this happened in Subsection 2.5.3), could also have carried away angular momentum.

SAQ 2.3

According to Table 2.1, FeS could not have condensed from the solar nebula unless the temperature had dropped to below about 700 K, so this value provides an upper limit on the possible temperature at which the chondrules originated.

SAQ 2.4

The lines to the left and right of the step in Figure 2.4 are parallel to the gas line, implying no change in composition throughout the solar nebula within each of these two regions. The most convenient place for us to read off the proportions of the three components is at the point where the column mass of dust steps up from just rock to rock plus ice.

(a) Immediately to the right of the step-up in column mass of dust, the column mass of rock plus ice is about 60 kg m^{-2} , whereas the column mass of gas at this point is about $2.8 \times 10^3 \text{ kg m}^{-2}$. To two significant figures, this value is also the same as the total column mass of the nebula. The amount of rock plus ice expressed as a fraction of this is given by:

$$\frac{60 \text{ kg m}^{-2}}{2.8 \times 10^3 \text{ kg m}^{-2}} = 0.021$$

[Comment: This result matches the 0.02 value for the relative abundance by mass of elements other than hydrogen and helium in the Sun (i.e. metallicity) given in Book 1, Subsection 1.5.3. Your reading of Figure 2.4 is likely to differ slightly from ours, because of difficulties in precisely reading values off the scale, but you should have got a similar result to ours. Note that the gas in the nebula will contain neon and other volatiles other than hydrogen and helium, and that conversely some hydrogen will occur in the dust, especially as H_2O , CH_4 (methane) and NH_3 (ammonia), so that you should not expect exact agreements between the value derived here and that from Book 1.]

(b) Immediately to the left of the step-up in column mass of dust, the column mass of rock is about 15 kg m^{-2} . The column mass of rock plus ice dust at the same distance (immediately to the right of the step-up) is 60 kg m^{-2} , so the ratio of rock to ice is given by:

$$\frac{\text{rock}}{(\text{rock plus ice}) - \text{rock}} = \frac{15 \text{ kg m}^{-2}}{60 \text{ kg m}^{-2} - 15 \text{ kg m}^{-2}} = \frac{15}{45} = \frac{1}{3}$$

[Comment: Your reading of the graph may have been different from ours, but you should have ended up with a fraction with a value of approximately 0.3. Note that this is actually an upper limit because Figure 2.4 shows only the contribution of H_2O to the ice, and does not include other volatiles such as ammonia that add to the amount of ice condensing at greater distances from the Sun. However, H_2O is considerably more abundant than any other ice-forming substance.]

SAQ 2.5

These are cumulative curves. As we follow a curve from right to left it climbs, which shows that the total (i.e. cumulative) number of bodies greater than the mass indicated on the horizontal axis is increasing. At times later than $t = 0$, the slope to the left of about 10^{16} kg shows that bodies of less mass than this have been produced. This can be explained by fragmentation during collision, even though collisions have also resulted in an increase in the maximum size of planetesimals (see bottom right of curves). [Comment: In the discussion of asteroid formation in Subsection 2.5.4 you saw that under certain conditions the rate of fragmentation can inhibit runaway growth entirely.]

SAQ 2.6

The **terrestrial** planets formed within 3 AU of the Sun, where the temperature was too high for water (the least volatile hydrogen-rich compound in the condensation sequence) to condense (as ice) from the solar nebula; most of their water probably arrived in the form of hydrated minerals. The terrestrial planets are too small to have collected much nebular gas directly by gravitation, and what they (or their planetary embryo precursors) did collect would have been lost during the Sun's T Tauri phase. The giant planets formed beyond 5 AU, so water was able to condense at Jupiter, and other hydrogen-rich substances such as methane and ammonia at greater distances. Thus planetesimals colliding to form the kernels of the giant planets would be rich in these materials as well as the more refractory elements. Furthermore, the kernels of the giant planets were sufficiently massive that each was able to capture and retain a large mass of gas (which would be mainly hydrogen and helium) directly from the nebula. This capture was interrupted by the onset of the T Tauri phase, which explains the decreasing volatile:rock ratio outwards from Jupiter (which began to form earliest) to Neptune (which began to form last).

SAQ 2.7

The large prograde satellites of Jupiter, Saturn and Uranus probably grew within protosatellite discs around each planet. Neptune's large retrograde satellite, Triton, may be a captured planetary embryo. Small satellites in prograde orbits can be formed by collisions involving larger ones, but those in retrograde orbits are more likely to be captured asteroids or comet nuclei.

SAQ 2.8

As ITQ 2.1 and the associated text discussed, the observable part of the Beta Pictoris dust cloud extends inwards to about 20 AU (where the orbit of Uranus would be). The dust suggests that the Beta Pictoris 'solar nebula' has probably completed its condensation phase, and grain growth by coagulation could be well advanced. There is no observational evidence bearing on whether planetesimal

formation and subsequent growth of planetary embryos has begun. However, the apparent absence of dust within about 20 AU of Beta Pictoris could mean planets have already formed in this region. Beta Pictoris is a main sequence star, and so should have undergone its T Tauri phase, but if so, the survival of dust is perplexing. Probably the best bet is that this is how our own Solar System would have looked while Uranus was growing but before the Neptune kernel had had chance to form. Subsection 2.5.3 suggests that there would be an interval of 20 million years between these events. [Comment: This is not to say that we should expect Beta Pictoris to form planets with the same masses and orbits as those in our own Solar System, because planetary formation must be influenced by the original mass and angular momentum of the nebula, and by subsequent random events (especially during and after runaway growth).]

SAQ answers and comments for Chapter 3

SAQ 3.1

(a) Accretional heating, especially as a result of giant impacts, makes it virtually certain that Venus is a differentiated body, in which case it must have a core.

(b) Its density shows that if its mantle is silicate then it must have a dense core (see ITQ 3.1), which is likely to be iron-rich, as in the Earth. This core is unlikely to be in vigorous motion, otherwise there would be a dipole magnetic field of a strength comparable with the Earth's. The most straightforward explanation for this is that Venus's core is entirely solid.

SAQ 3.2

Our attempt is shown in Figure 3.15. We do not expect yours to look exactly the same. The main point is that the temperature just below the surface should be hotter than the 9×10^6 year example in Figure 3.6, because there has been less time for accretional heat to be radiated to space.

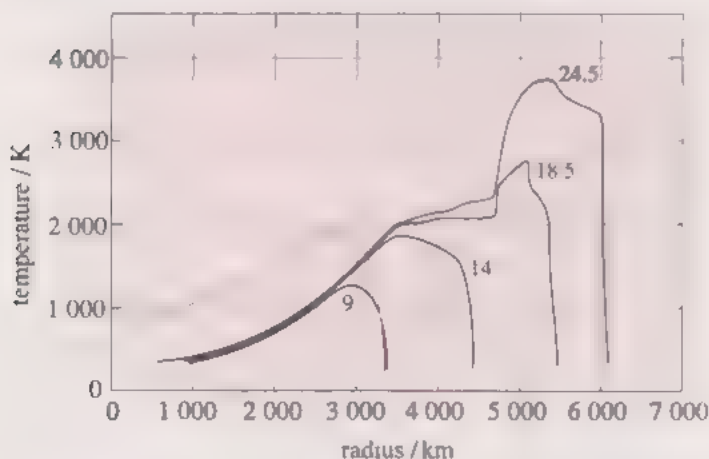


Figure 3.15 Temperature against depth for a 3400 km radius embryo (red line); the answer to SAQ 3.2.

SAQ 3.3

Giant impacts are an inevitable stage of the growth of planets from planetary embryos outlined in Subsections 2.5.2 and 3.3.2. In the specific case of the Earth–Moon system, as discussed in Section 3.4, the Moon's anomalously low core volume is compatible with its formation from debris from the mantle of a planetary embryo that had collided with the proto-Earth, and whose core sank through the Earth's mantle to add to the Earth's core. The Moon's enrichment in refractory elements and depletion in volatiles relative to the Earth can be explained by the same process. So can the particularly low abundances of siderophile elements in the Moon's mantle and crust, which would have already been depleted in these elements before they were further scavenged by iron during the differentiation event that produced the Moon's tiny core. Similarities in the proportions of the stable isotopes of oxygen in the Earth and Moon argue that these two bodies are formed from matter that condensed in the same region of the solar nebula (which would make the right sort of collision easier). Any slight differences in oxygen isotopes would have been blurred by mixing of their mantles during the collision.

SAQ 3.4

To two significant figures, the ratio between the periods of Enceladus and Mimas is 1.5 (i.e. 3 : 2), the ratio between Dione and Enceladus is 2.0 (i.e. 2 : 1), and the ratio between Mimas and Dione is 2.9 : 1 (which is not quite 3 : 1). This means that Mimas and Enceladus, and Enceladus and Dione, are in orbital resonance, but there is no orbital resonance between Mimas and Dione. Orbital resonance means inner satellites overtake outer ones at the same positions in their orbits, and the repeated gravitational tug between them prevents their orbits from becoming circular. Their tidal bulges are flexed, and this leads to internal heating. [Comment: Of these three satellites, Enceladus

shows plenty of signs of recent tidal heating, whereas Mimas looks to have been dead for billions of years.]

SAQ 3.5

(a) The lithosphere must thicken with age, as the rate of radiogenic heat production declines.

(b) No. Tidal heating does not decrease exponentially with age, like radiogenic heating. In systems involving several satellites, a satellite can experience several interludes of orbital resonance (video sequence 5), each of which would lead to a pulse of heating. In such a situation the lithosphere would thicken and thin with time, as the rate of heating fluctuated.

SAQ 3.6

The Earth's core, crust and mantle are all defined seismically, on the basis of changes in seismic speed, which

reflect differences in composition. The core is iron-rich whereas the mantle and crust are formed of silicates. The crust is less dense and richer in SiO_2 than the mantle. Interfaces between these three units are marked by sharp changes in the speed at which seismic waves are transmitted. On the other hand, the lithosphere is distinguished from the asthenosphere on the basis of strength, and the associated differences in the way in which heat is transferred. The lithosphere is rigid and comprises the crust and the uppermost mantle. Heat passes through the lithosphere by conduction and advection, though about two-thirds of the global transfer of heat from the interior to the surface is accomplished by plate recycling involving wholesale creation and destruction of the lithosphere (Figure 3.14). The asthenosphere is effectively the whole of the mantle below the lithosphere; it is weaker, and transfers heat mainly by convection.

SAQ answers and comments for Chapter 4

SAQ 4.1

Because they are clustered, rather than randomly scattered, and show 'herringbone' features, most of the small craters are probably secondaries, including those forming the prominent chain in the middle distance. Not all are from Copernicus, however. In the left half of the photograph, many 'herringbone' secondary impact features are aligned northwest-southeastwards. They actually originate from the crater Eratosthenes, just off the photograph at top left. [Comment: The photograph is an oblique view taken by an Apollo astronaut. A more conventional view of Copernicus appears in Figures 4.4.]

SAQ 4.2

There is no simple answer to this question, but it must relate to the timing of impact relative to the rate of lithospheric thickening as the Moon and Callisto cooled. The Orientale basin formed relatively late, when the Moon's (silicate) lithosphere was thick and strong enough to support substantial topographic features, but Valhalla presumably formed in a thinner (icy) lithosphere.

SAQ 4.3

All the innumerable craters are perfectly circular, and there is no suggestion of any evidence of 'herringbone' patterns. Thus, the photographic evidence suggests they are of primary impact origin. The largest craters visible are deep, bowl-shaped depressions, with no suggestion of internal structures or terraced walls. Thus, from the morphological criteria in Subsection 4.3.2, they cannot be more than about 15 km in diameter. [Comment: Measurements show that the pair of prominent craters at left centre are only about 3 km in diameter.]

SAQ 4.4

(a) Many of the clusters of craters are most likely secondary craters, derived from some major impact site beyond the

frame of the photograph, probably towards the top left. Compare with the secondary craters in Figure 4.6. (b) The central strip, roughly symmetrical around the prominent fissure, has considerably fewer craters. It appears to have been resurfaced at some time, perhaps by lavas welling up from the fissure. (c) It would not be easy to make good estimates of the absolute ages of the surfaces, because all of the craters are small – only a few hundred metres in size – and therefore many of them are most probably secondaries.

SAQ 4.5

This question is readily answered by reading off values from the two curves – the left-hand scale is easier to use. For example, on Earth a 2 billion year old surface would have a crater density of about $7.3 \times 10^{-5} \text{ km}^{-2}$ (Figure 4.21), whereas a 2 billion year old surface on the Moon has a crater density of about $5.5 \times 10^{-5} \text{ km}^{-2}$ (Figure 4.17). Thus, the production rate on Earth must be about $7.3/5.5$ of that of the Moon, or about 1.3 times greater. Although the Earth and Moon are close together in space, Earth's crater production rate has been higher because its greater mass means that it can 'capture' gravitationally more potential impactors. Given the broad spread of the uncertainties suggested by the plots, the value of 1.3 is at best very approximate.

SAQ 4.6

The plot would show a perfectly ordinary gradient – although the population of craters would initially have been skewed towards small sizes, the fact that this was followed by 'ordinary' bombardment until saturation would mean that evidence of the earlier episode would now be completely obliterated. [Comment: If the situation were reversed (ordinary population followed by small) this would be obvious, resulting in a steeper plot – more small craters.]

SAQ 4.7

The smallest craters with central peaks are about 20 km across. Following the arguments in Subsection 4.3.2, one would expect the smallest craters with central peaks on Mercury to be smaller than on the Moon (Table 4.1), that is

less than about 15 km in diameter. However, the image covers only a small, possibly not representative area of the surface. Furthermore, the fact that there are many craters much larger than 20 km that lack central peaks shows that there is only a loose relationship between crater structure and size.

SAQ answers and comments for Chapter 5

SAQ 5.1

In detail, the physics of tidal dissipation and its effects on the orbital evolution of the Moon are complex, but there should be a little heat generated by tidal dissipation in the Moon. The Moon's orbit around the Earth is slightly elliptical, hence dissipation of tidal energy results from the oscillation of the tidally induced bulge about a mean position. If the Moon's orbit were circular, the bulge would be stationary (video sequence 5).

SAQ 5.2

The temperature in the interior of a large icy body would be controlled only by relicts of any primordial heat sources. These would be primarily from energy liberated by impacts during accretion, and gravitational self-compression of the growing body. There would be no heat from radioactive isotopes or core formation, or from tidal dissipation *unless* the planet had a massive satellite.

SAQ 5.3

If a planet exhibits any plate recycling, it must have a mobile, convecting asthenosphere, and a rigid lithosphere. Assume that there is insufficient tidal heating to prevent cooling. As the planet cools, the thickness of the lithosphere will gradually increase, making plate recycling more difficult, and so surface volcanism (advection) at a few sites may become more important. Plate recycling will fade away entirely, and so ultimately will surface volcanism, leaving conduction as the only heat loss mechanism. Although the *proportion* of heat transferred by conduction increases, the *rate* of heat transfer decreases through time. A summary triangular diagram would resemble Figure 5.17 with the path terminating at the 'conduction through the lithosphere' vertex. [Comment: The position of the Earth today is shown for reference.]

SAQ 5.4

Because Venus is closely similar in size, mass and composition to the Earth, one would expect its basalts to resemble the commonest basalts on Earth, MORB lavas. [Comment: The crude analyses performed by spacecraft that have landed on Venus show that the lavas are indeed similar to MORB. However, they have quite high MgO contents, almost the same as the values found in terrestrial komatiites. TV programme 4, *Venus unveiled*, tells you more about this.]

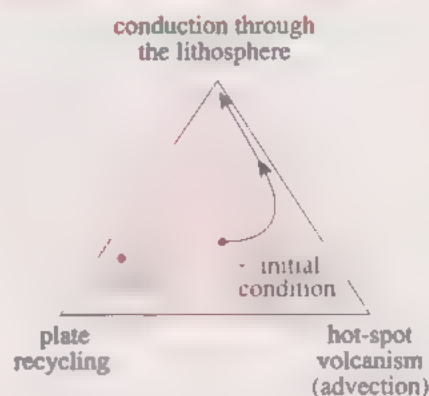


Figure 5.17 Triangular diagram showing the heat loss path for a cooling planet. (For SAQ 5.3 answer.)

SAQ 5.5

This question can be answered without much calculation by spotting that all of the variables in Equation 5.3 remain substantially unchanged, except for the atmospheric density β . From Table 1.1a, we see that, at the surface, Venus's atmosphere is about 55 times as dense as that at the Earth's surface. Thus, the terminal fall speed on Venus must be $1/\sqrt{55}$ that on Earth, or $100 \text{ m s}^{-1} \times 1/\sqrt{55} = 13.5 \text{ m s}^{-1}$. [Comment: If you allow for Venus's slightly lower surface gravity, 8.9 m s^{-2} instead of 9.8 m s^{-2} for Earth, you will find that the actual terminal fall speed is slightly lower than this, about 12.9 m s^{-1} .]

SAQ 5.6

Because the volatile content is high (1.25%) the eruption would be an explosive one, producing pyroclastic rocks. But because the Moon lacks an atmosphere altogether, the ejected material would be distributed ballistically.

SAQ 5.7

Equation 5.3 shows that terminal fall speeds are proportional to the square root of surface gravity, and inversely proportional to atmospheric density. Mars's surface gravity is less than half the Earth's (3.7 m s^{-2} versus 9.8 m s^{-2} ; Table 1.1a), tending to reduce terminal fall speeds. But because Mars's atmosphere is so much more tenuous (its atmospheric density is about one-seventieth that of the Earth's), terminal fall speeds for pyroclastic materials are actually much *higher* on Mars than on Earth. Thus, they would spend less time falling through the atmosphere, and the isopleth for a given size fragment would be more nearly circular than on Earth, and (other things being equal) cover a smaller area. At smaller

fragment sizes, the terminal speeds are smaller (Equation 5.3), so the differences in isopleth shape would diminish, though again the Martian ones would cover a smaller area.

SAQ 5.8

Only (vii) carbon dioxide. This gas has a sufficiently high molecular mass (Section 1.4, Figure 1.10) to be retained, and is also abundant enough to be a plausible component of Io's atmosphere. The others are too lightweight, except for carbon disulphide (viii), which is rare in planetary atmospheres (Table 1.1a). [*Comment:* Carbon disulphide is

a disgustingly smelly gas, with an odour like rotten cabbage.]

SAQ 5.9

Io's heat production is likely to lead to high internal temperatures, and high degrees of partial melting. As discussed in Section 5.2, this means that its lavas are likely to be 'primitive', i.e. derived directly from (nearly) chondritic compositions. They may resemble terrestrial komatites.

SAQ answers and comments for Chapter 6

SAQ 6.1

Mass spectrometry identifies molecules on the basis of their masses, and so (except in instruments with very high resolution) would not distinguish between $^{14}\text{N}_2$ (RMM = 28) and $^{12}\text{C}_2^{1}\text{H}_4$ (RMM = 28). [*Comment:* It would, however, enable us to distinguish between isotopic variants of molecules, e.g. $^{12}\text{C}^{16}\text{O}$ and $^{13}\text{C}^{16}\text{O}$.]

Each of these chemically distinct molecules could be identified by gas chromatography, provided that the instrument had been calibrated for each substance. [*Comment:* However, no information about isotopes is usually obtained by gas chromatography.]

SAQ 6.2

Diatomic molecules without electric dipoles cannot be detected by infrared spectroscopy, and so H_2 would not be detected. The heteronuclear molecules CO and HCl would be detected.

SAQ 6.3

By comparison with the terrestrial planets, a high albedo indicates a highly reflecting surface. This suggests either cloud cover or some nearly white surface. [*Comment:* The atmosphere of Triton is very thin and essentially cloud-free; the surface is composed of highly reflective, icy material.] A high albedo and large distance from the Sun (so low solar flux density) are consistent with a low temperature.

SAQ 6.4

By absorbing infrared radiation that is not otherwise absorbed in the Earth's atmosphere, the chlorofluorocarbons add to the greenhouse effect. The temperature of the Earth's troposphere will therefore be increased by their presence. [*Comment:* This is in addition to, and quite separate from, the role of CFCs in depleting the ozone layer, which you will meet in Subsection 6.5.3.]

SAQ 6.5

At the average surface temperature of Mars, about 223 K, a vertical line on Figure 6.22 intersects the curve very close to the solid/liquid transition at point T. However, the partial pressure of CO_2 at the surface of Mars (about 0.006 bar – see Table 1.1a – essentially the total surface pressure)

suggests that CO_2 will not condense. [*Comment:* Near the poles during winter, when temperatures drop below 150 K, CO_2 will condense as solid according to Figure 6.22.]

SAQ 6.6

The initial speed of this piece of atmosphere at the equator will be the planet's surface rotation speed, that is $0.997/243$ of that of the Earth. So the gain in speed relative to the surface will also be $0.997/243$ of that of the gain on the Earth, which was calculated in Subsection 6.4.4 to be 134 m s^{-1} . This is $0.997 \times 134 \text{ m s}^{-1}/243 = 0.55 \text{ m s}^{-1}$. As Venus rotates in a retrograde direction, the direction will be east to west. [*Comment:* 0.997 days is the Earth's sidereal rotation period.]

SAQ 6.7

Addition of the three equations leads to the following result:



[*Comment:* This scheme is part of the chemistry that accounts for the destruction of ozone, which involves atomic oxygen, as discussed in Subsection 6.5.3.]

SAQ 6.8

(a) In general, exothermic reactions in which bonds are formed but not broken, or those involving bond formation and breaking in which only a diatomic molecule is produced, require the presence of a third body M. The third body acts to remove the potential energy released in the reaction. In Reaction 6.35, the tetra-atomic molecule NO_3 is sufficiently large to store this energy in its vibrations, so the third body M is not required.

(b) Reaction 6.36 involves bond breaking but not bond formation. Energy is required to break chemical bonds, and so this reaction is deduced to be endothermic.

(c) The hydroxyl radical OH is the general oxidizing substance for hydrocarbons (such as CH_4). It operates both day and night. In contrast, NO_3 is photodissociated by UV radiation, so its concentration during daylight is very low and it is unable to compete with OH as an oxidizing agent. Only at night is its lifetime sufficiently long to enable it to compete.

SAQ 6.9

The Earth's atmosphere is oxidizing because it contains O_2 , which is at least capable of keeping Fe oxidized to the Fe^{3+} state. On Mars, oxidizing substances are produced in a variety of ways, for example hydroperoxyl (HO_2) from the photodissociation of water. These, too, ensure that iron is present as Fe^{3+} , and indeed the metal peroxides contained in the surface soil are themselves oxidizing.

On Venus, photodissociation is limited to the atmosphere above cloud level, so the surface is not exposed to molecules more oxidizing than CO_2 . [Comment: Note, however, that there are suggestions that in spite of the high temperatures in the lower atmosphere, some H_2SO_4 might reach the surface as 'rain'. This substance has some oxidizing ability and there is some evidence of Fe^{3+} on the surface of Venus. See also Subsection 6.6.5.]

SAQ 6.10

The atmosphere of Mercury is the outcome of the balance between the capture of solar wind and thermal escape. As Equations 6.38 and 6.39 show, the escape is much more likely for lighter molecules and atoms. An atom of He is twice as massive as the H_2 molecule, so H_2 escapes faster than does He. Also, H_2 is prone to photodissociation to produce atomic hydrogen, which escapes even more readily.

SAQ 6.11

These atmospheres are the result of outgassing. If the non-volatile components were similar, the amounts of ^{40}K from which ^{40}Ar is produced by decay would be similar. So the relative amounts of ^{40}Ar reflect the relative extents of outgassing. This is particularly low for Mars.

SAQ 6.12

Evidence of former oceans would be sought in the sedimentary rocks that would have formed as carbonates from the CO_2 that would have been dissolved in the oceans. Proof of this would require analysis of the surface rocks for carbonate deposits, a formidable task. [Comment: Indeed, carbonates are likely to be unstable at the prevailing temperatures, so that such evidence may not exist today even if there had been oceans.] Other evidence would be the recognition of ancient shorelines on radar images of Venus, of which there is at present no evidence (see TV programme 4).

SAQ 6.13

The initial effect of an added greenhouse gas is to raise the atmospheric temperature. This would cause the evaporation of some CO_2 and H_2O from the polar caps and permafrost. With their added contribution to the greenhouse effect, further evaporation would occur, although a 'runaway' greenhouse effect might not be achieved. If the temperature and pressure rose sufficiently, water could exist in the liquid state, the basis for the far-fetched suggestion in the question.

SAQ 6.14

The source of O_2 in the Earth's atmosphere, apart from traces produced by H_2O photodissociation, is photosynthesis, by which process O_2 is produced from CO_2 . In the early atmosphere, the planetary surface would have been exposed to the life-damaging UV radiation that is presently absorbed by ozone in the stratosphere. At this time, the protection of life from this radiation was provided by liquid water. [Comment: It has been estimated that the earliest forms of life would have required a 10 m depth of water to avoid damage by this radiation.]

SAQ answers and comments for Chapter 7

SAQ 7.1

The far greater mass of Jupiter means that self-compression is greater and that, were Jupiter and Rhea made of the same materials, the density of Jupiter would be greater. Having the same average density is thus evidence *against* the two being made of similar material and *for* them being different.

SAQ 7.2

Mercury has a very high density, and a substantial layer (which would be needed to produce the magnetic dipole) would add to the average density, making it far higher than the observed value.

Mercury is not a very abundant element and so a large amount of it would be ruled out if we assume that the relative abundances are roughly solar. [Comment: The form of mercury at the temperatures and pressures in the interior of Saturn is not given, but at such high pressures it is likely that mercury will still be a conducting fluid.]

SAQ 7.3

Both Uranus and Neptune are observed to have magnetic fields, but the electrically conducting liquid layer is found to lie not in the core but farther out. The observed magnetic field data thus do not rule out the model but do not support it as an explanation of the magnetic fields in these planets.

Neptune's internal heat shows that its core is hot, but is not necessarily molten rock. There is no observational evidence for a hot interior in Uranus, so on the face of it this does not fit the model.

SAQ 7.4

(a) The proposed model must fit the data for the average density and the variation of density with depth (both inferred from gravitational field measurements) and the observed magnetic fields and (for Neptune) the excess heat. Rocky cores are not essential to produce the average density: as long as it can be shown that the proposed

interior is sufficiently dense to produce the average density of the planet, then there is no conflict here. Gravitational field measurements indicate a dense core, but this is not necessarily rocky. The magnetic fields originate in layers outside the core, so this evidence does not rule out an icy core. The interior heat is not well understood, but does not necessarily conflict with the model.

(b) The theory of planetary formation outlined in Chapter 2 assumes a large planetary kernel of rocky and icy material, which then sweeps up surrounding material. A wholly icy model poses problems for explaining how Uranus and Neptune began to form.

SAQ 7.5

This may be evidence against the theory of heat excess, as helium from the molecular layer would have to replace that lost from the metallic layer by differentiation. However, helium in the molecular layer would in turn be replaced by helium from the atmosphere. Thus it is possible that only enough helium has separated out for depletion to show up in the atmosphere. [Comment: Saturn's atmosphere does indeed seem to be depleted in helium – see Section 7.3.]

SAQ 7.6

Neon, like helium, is a noble gas and forms no compounds. It cannot therefore be detected through vibrational spectroscopy. Its atomic spectrum would have to be used. [Comment: This is more accessible than that of helium.] The solar abundance of neon is much less than that of helium (Book 1, Chapter 2 Appendix) and so we would expect only a small amount to be present in the atmosphere. This would make it harder to detect by its atomic spectrum and rules out the indirect methods used for helium.

SAQ 7.7

The most likely oxygen-containing molecule is water, H_2O , and this condenses at much higher temperatures than ammonia or methane. Most of the atmospheric oxygen could be frozen out as clouds below the observed atmosphere, leaving only a trace below the detectable limit. There must also be a considerable amount of oxygen locked up as H_2O in the icy part of the planet beneath the atmosphere.

SAQ 7.8

Nitrogen is a homonuclear molecule N_2 , and so could not be observed through its vibrational or rotational spectrum (Section 6.2). Much ultraviolet radiation is absorbed by the Earth's atmosphere, so that it would be very hard to observe its electronic spectrum from the ground.

SAQ 7.9

Uranus and Neptune are much colder than Jupiter and Saturn. Ammonia will condense out at the temperatures found on Jupiter and Saturn, but methane needs the lower temperatures found on Uranus and Neptune. [Comment: This applies at the observed partial pressures. If the abundance of methane were much higher then it would be

possible for it to condense even at the temperature found on Jupiter – see Box 6.1.]

SAQ 7.10

The wind velocities are measured relative to the rotation velocity of the magnetic field. It is assumed that the layer where the field originates rotates at the same velocity as (much of) the planetary interior. This is important because the wind velocities are held to be relative to the planetary interior. [Comment: The pressures are so high in the interior that the materials will be viscous and so this assumption is not unreasonable.]

SAQ 7.11

A negative wind velocity corresponds to a piece of atmosphere rotating more slowly than the magnetic field, albeit in the same direction. [Comment: A similar effect occurs if you are travelling in a fast train and pass a slower-moving train. If you look out at the slower train it appears to be going backwards.]

SAQ 7.12

Jupiter has an internal heat source and also receives energy from the Sun.

The rate of decline of temperature with height in the troposphere is close to that expected for heat transport by convection.

The rate of loss of heat to space is low owing to the low temperatures of the radiating levels in the atmosphere.

SAQ 7.13

Hurricanes die out because they lose energy. On Earth, energy is lost when the wind interacts with surface features. On Jupiter, there is no surface. The other method of energy loss is emission of heat (infrared radiation). Less energy is lost this way on Jupiter because the atmospheric temperatures of Jupiter at the relevant levels are lower. The rate of loss of energy is thus lower on Jupiter and the storms persist far longer.

SAQ 7.14

An internal heat source would be expected to heat all latitudes uniformly. The explanation given assumes that the convection cells on Neptune are driven by solar radiation, which warms mid-latitudes more than the poles or the equator. [Comment: At the distance of Neptune from the Sun, heating from internal energy could well be more important than solar heating.]

SAQ 7.15

Yes, ethane and ethyne (C_2H_6 and C_2H_2) have been observed in Neptune's atmosphere (Table 7.1).

SAQ 7.16

H_2O condenses out at higher temperatures than ammonia or methane and so little is left in the regions of the atmospheres where photochemical reactions occur.

SAQ answers and comments for Chapter 8

SAQ 8.1

It is an impact-produced sample of some description. It is probable that the chondritic, achondritic and metal-rich parts of the meteorite were all formed separately by normal processes. These were then assembled at the surface of an asteroidal body during an impact event – for instance, an iron meteorite may have impacted a silicate-rich parent body. [Comment: The meteorite described is not a hypothetical example but one member of a well documented, albeit rather rare, class of unusual samples.]

SAQ 8.2

In three cases, data from photographic networks have been used to compute the orbits of meteorites, which can be traced back to the asteroid belt. Furthermore, photometry has been used to classify asteroids into various different types, which can be related to known meteorite groups (e.g. carbonaceous chondrites are thought to arise from C-type asteroids).

SAQ 8.3

The surface area of the Earth = $4\pi \times (6\,378\text{ km})^2 = 5.11 \times 10^8\text{ km}^2$.

If we assume (reasonably) that meteorites fall evenly over all parts of the Earth, then in one year a single meteorite with a diameter greater than 0.1 m falls in an area of

$$\frac{5.11 \times 10^8}{20\,000}\text{ km}^2 \approx 25\,600\text{ km}^2$$

In an area of $25\,600\text{ km}^2$ there could be up to $\frac{25\,600}{50} = 512$ villages each occupying 50 km^2 , so on average, a meteorite of $>0.1\text{ kg}$ mass would fall in a village covering 50 km^2 less than about once every 500 years.

SAQ 8.4

Compared with C1 meteorites, C3 samples are enriched in Ca and Al. This can be explained by a relatively higher proportion of Ca- and Al-rich minerals in C3 meteorites. A logical conclusion here would be that C3 meteorites are the most enriched in CAIs. [Comment: This is indeed true.]

SAQ 8.5

That the data from the matrix materials plot on a line of slope 0.5 suggests that the oxygen isotopes in these minerals have been subjected to processes of isotope fractionation, which have resulted in $\delta^{18}\text{O}$ being twice $\delta^{17}\text{O}$. This is evidence that the matrix materials were formed on a single parent body, rather than in the solar nebula. Since the line of slope 0.5 is not coincident with the TFL, it shows that the reservoir from which the carbonaceous chondrites were extracted was not exactly the same as that which formed the Earth.

SAQ 8.6

A very simple way would be to measure the stable isotope composition of carbon in each form. If there is no difference in $^{12}\text{C}/^{13}\text{C}$ ratios then the two forms might be related. [Comment: In reality, there is a distinct difference in carbon isotope composition between the diamond and the amorphous carbon. Although it is entirely possible that diamond can be formed from amorphous carbon by shock processes (which result in high pressures), the isotope evidence is at variance with such an origin. Several other lines of evidence also mitigate against shock, so the two forms of carbon appear to come from entirely different sources.]

SAQ 8.7

If we assume that all meteorite parent bodies started off with a chemical composition like the C1 carbonaceous chondrites [Comment: rather like the whole Earth, see Section 5.2] then in order to obtain a metal-rich body there has to be extensive melting and differentiation. Upon total melting, silicate materials migrate to the outermost parts of an asteroid while metal sinks to the centre of the body. If the metallic asteroids we can observe today are original cores then the silicate crusts and mantles of these bodies must have been removed (most probably by collisional processes in the asteroid belt). [Comment: Unlike the asteroid Vesta, which is still largely intact, a metallic asteroid represents the remnants of an original parent body.]

SAQ 8.8

With just the knowledge that Kodaikanal has a Rb–Sr age of 3.8 billion years, we may be inclined to think that this was the age of formation. However, as stated in the text, most meteorite ages fall in the range 4.45–4.53 billion years. The impact-produced rocks from the Moon document a period of Solar System history when planetary bodies were undergoing intense bombardment by meteoroids. The age of Kodaikanal could record an impact event during the same era, but in this case it occurred in the asteroid belt. If such an event caused localized melting of the parent body, the Rb–Sr age would document the timing of that event. [Comment: Kodaikanal is demonstrably a meteorite, being composed predominantly of iron–nickel metal, and having a Widmanstätten pattern. There is no possibility that Kodaikanal represents a piece of ejected lunar material.]

SAQ 8.9

Diagnostic features of an SNC meteorite are as follows.

- It may have a fusion crust, produced when the meteorite descended through the atmosphere.
- The actual fall may have been witnessed.
- There may be evidence of an impact crater.
- It would have an oxygen isotope composition different from that of a terrestrial rock.

SAQ 8.10

First and foremost, it would be advantageous to determine the crystallization age of the sample using a radiometric dating system, such as Rb–Sr. If the sample turned out to be about 4.5 billion years old then you would have good grounds for thinking it was a meteorite like Juvinas. If the age was considerably younger than this (i.e. ≤ 1.3 billion years old) you could be reasonably certain it was an SNC meteorite.

SAQ 8.11

(a) Nitrates formed at the present day would have $^{14}\text{N}/^{15}\text{N}$ ratios similar to atmospheric nitrogen (i.e. 165, as measured by the Viking landers). (b) Similar minerals formed in the past would have $^{14}\text{N}/^{15}\text{N}$ ratios that reflected the extent of atmospheric evolution at the time of formation (i.e. any value between 165 and 272, which is the value that Mars is assumed to have started with). [*Comment:* Consider how important measurements of this nature would be if samples could be obtained from known places on the Martian surface, for instance by a space mission. The $^{14}\text{N}/^{15}\text{N}$ ratios of nitrate minerals could possibly be used to date the materials, and by extension the formation regions.]

SAQ 8.12

It may be anticipated that collisional processes in the asteroid belt will produce molten materials. These may be distinguished from cometary dust, which retains its primitive, unmelted, and in some cases porous, character. In addition, we may expect asteroidal debris to be more like meteoritic material than the cometary dust which, although approximately chondritic, does not match any of the well-known meteorite groups. [*Comment:* Most research carried out so far has concentrated on the porous cometary dust grains. Thus, we do not have much information about particles that may be asteroidal in origin.]

SAQ 8.13

As ever, oxygen isotope measurements would show whether or not the samples fall on the TFL (Subsection 8.3.2). If they fall on a parallel line, or do not fall on a line at all, a non-terrestrial origin is indicated. In addition, detection of isotopes produced by cosmic ray interaction (see Subsection 8.5.3) would demonstrate that micrometeorites had originated from outside the Earth's atmosphere.

Acknowledgements

We are grateful to the two assessors for this book, Pamela Tonks and Lionel Wilson, whose perceptive comments on any early draft were an important guide to us in preparing the final thing.

Grateful acknowledgement is also made to the following sources for permission to reproduce material in this book:

Chapter 1

Figure 1.5a, b Redrawn from *Dance of the Planets* graphics software, ARC Inc.

Chapter 2

Figures 2.5, 2.6, 2.7, 2.8 © Copyright Harold Weaver and George Weatherill/Space Telescope Science Institute; Figure 2.9 Jet Propulsion Laboratory, Pasadena, California, image no. P-41525; Figures 2.11, 2.12 Dr Bradford A. Smith, NASA Goddard Space Flight Center, World Data Center A for Rockets and Satellites.

Chapter 3

Figure 3.6 W. M. Kaula (1974), 'Thermal evolution of Earth and Moon growing by planetesimal impact', *Journal of Geophysical Research*, vol. 84, no. B3, American Geophysical Union; Figure 3.7 W. Benz *et al* (1987), 'The origin of the Moon and the single-impact hypothesis II', *Icarus*, vol. 71, Academic Press Inc.; Figure 3.8 T. Matsui and Y. Abe (1986), 'Origin of the Moon and its early thermal evolution', in Hartmann, W. K., Phillips, R. J. and Taylor G. J. (eds), *Origin of the Moon*, Lunar and Planetary Institute, Houston; Figure 3.9 A. G. W. Cameron and W. Benz (1991), 'The origin of the Moon and the single-impact hypothesis IV', *Icarus*, vol. 92, Academic Press Inc.

Chapter 4

Figures 4.1, 4.2, 4.4, 4.5, 4.6, 4.7, 4.9, 4.10, 4.11, 4.12, 4.13, 4.14, 4.19, 4.20 NASA Goddard Space Flight Center, World Data Center A for Rockets and Satellites; Figure 4.3a, b Courtesy of Dr Peter W. Francis; Figure 4.8 Institute of Geological Sciences.

Chapter 5

Figure 5.1a, b, 5.9–5.16 NASA Goddard Space Flight Center, World Data Center A for Rockets and Satellites; Figure 5.4 L. Lirer, T. Pescatore, B. Booth and G. P. Walker (1973), 'Two plinian pumice-fall deposits from Somma-Vesuvius, Italy', *Geological Society of America Bulletin*, vol. 84, Jan–Apr 1973, Geological Society of America; Figure 5.7 L. Wilson, R. S. J. Sparks and G. P. L. Walker (1980), 'Explosive volcanic eruptions IV: The control of magma properties and conduit geometry on eruption column behaviour', *Geophysical Journal Royal Astronomical Society*, vol. 63, Blackwell Scientific Publications Ltd; Figure 5.8 Courtesy of Dr Peter Francis.

Chapter 6

Figure 6.1 Courtesy of Fleetway Editions Ltd; Figure 6.25 J. C. Goedheer (1966), in Vernon, L. P. and Seely, G. R. (eds), *The Chlorophylls (Physical, Chemical and Biological Properties)*, Academic Press Inc.; Figure 6.31 NASA Goddard Space Flight Center, World Data Center A for Rockets and Satellites; Figure 6.32a Dr Phillip Playford, Geological Survey, Dept of Mines, Western Australia; Figure 6.32b M. R. Walter, MacQuarie University, New South Wales, Australia; Figures 6.33, 6.34 Adapted from R. P. Wayne (1991), *Chemistry of Atmospheres*, by permission of Oxford University Press.

Chapter 7

Figure 7.16 European Southern Observatory; *Figure 7.18* R. F. Beebe (1990), 'Queen of the giant storms – Jupiter's red spot and Neptune's dark spot', *Sky and Telescope*, vol. 80, no. 4, © 1990 Sky Publishing Corp., reproduced with permission.

Chapter 8

Figure 8.1 U. B. Martin (1983), 'Extraterrestrials have landed on Antarctica', *New Scientist*, 17 March 1983, IPC Magazines Ltd; *Figure 8.2* From *Meteorites Their Record of Early Solar-System History* by John T. Wasson, copyright © 1985, by John T. Wasson, reprinted with permission of W. H. Freeman and Company; *Figure 8.6* Courtesy of Xia Gao, McDonnell Center for Space Sciences, University of Washington; *Figure 8.9* H. J. Melosh (1984), 'Impact ejection, spalation and the origin of meteorites', *Icarus*, vol. 59, Academic Press Inc.; *Figure 8.10* Reprinted with permission from *Nature*, R. O. Pepin 'Evidence of martian origins', vol. 317, 10 October 1985, copyright © Macmillan Magazines Ltd; *Figure 8.12* R. M. Walker, McDonnell Center for the Space Sciences; *Figure 8.13* Adapted from J. P. Bradley, S. A. Sandford and R. M. Walker, (1988), 'Interplanetary dust particles', in Kerridge, J. F. and Mathews, M. S. (eds), *Meteorites and the Early Solar System*, The University of Arizona Press, © copyright 1988; *Figure 8.14* Adapted from S. A. Sandford and R. M. Walker (1985), 'Laboratory infrared transmission spectra of individual interplanetary dust particles from 2.5 to 25 microns', *The Astrophysical Journal*, 291, © 1985, The American Astronomical Society, all rights reserved.

Index

- ²⁶Al – see aluminium-26
- ablation 207
- abundances of elements 31–32, 48
- accretion 38, 42
- accretionary heating 53, 54, 59, 222, 251, 253
- achondrites 206, 220
- advection 65
- albedo 126, 128, 134, 254
- ALH A81005 (meteorite) 226
- Allende (meteorite) 210–212, 215–216, 218
- aluminium-26 54, 60, 222–223
- Amalthea 45
- amino acids 216–217
- ammonia (NH₃) 45, 97, 173, 178, 179, 184–185, 196–197, 201, 256
- ammonium hydrogen sulphide (NH₄HS) 185
- angular momentum 29–31, 62
- Antarctica 204, 225
- Apollo Moon landings 56, 70, 84
- argon 158, 255
- asteroid belt 16–18, 257
- asteroids 10, 16–19, 42, 45, 207, 221, 222, 234, 257
- asthenosphere (of a planetary body) 63–64, 77, 78, 80, 252
- astronomical unit (AU) 11
- atmosphere 41, 55, 230–231
- atmospheric structure 127–134, 182–184, 186–187

- basalt 96, 109, 113, 221, 226
- belts 189, 191
- Beta Pictoris 27, 34, 35, 251
- bipolar outflow 31
- buoyancy (of magma) 99

- ¹²C – see carbon isotopes
- C-type asteroids 19, 209
- C1 carbonaceous chondrites 95, 210, 257
- calcium–aluminium-rich inclusions (CAIs) 212, 213–215, 218, 223, 234, 257
- calderas (volcanic) 105
- Callisto 16, 44, 252
- Cape York (meteorite) 220–221, 223–224
- carbon 34, 145, 216, 236, 257
- carbon cycle (on Earth) 146
- carbon dioxide (CO₂) 126–127, 139, 155, 254, 255
- carbon isotopes 123–124, 236, 257
- carbonaceous chondrites (see also C1 carbonaceous chondrites) 31, 32, 48, 95, 158, 206, 210–211, 218, 230, 234, 257
- carbonate deposits 127, 155, 159, 216, 255
- Cassini's division 18
- central peaks 80
- Ceres 16
- CFCs – see chlorofluorocarbons
- CH₄ – see methane
- Chapman scheme 147–149
- Charon 7
- chassignites (see also SNC meteorites) 225
- chemical compounds 33, 34
- chemical energy 140–141
- chemical equilibrium 139, 141–142, 145, 149, 184
- chemical reactions 139–143
- Chicxulub crater 69
- Chiron 24, 249
- chlorofluorocarbons (CFCs) 138, 148, 165, 254
- chondrites (see also carbonaceous chondrites) 37, 221
- chondritic composition 32, 96
- chondritic Earth model 95
- chondrules 37, 38, 205–206, 212, 215, 234, 250
- circumstellar discs 27
- circumstellar dust clouds 34
- clathrate compound 152, 181
- cloud 124, 126–127, 149, 178, 183, 187
- cloud colour 185
- cloud formation 126, 133–134, 184–185
- CNO cycle 213
- coagulation 36, 37, 38, 52
- collision rate 141
- collisions 35–36, 38, 39, 42, 45
- column mass 36, 126

- coma (of a comet) 21
- comet Halley 20, 24, 233, 249
- comets 21, 42, 45, 209, 226, 234
- compatible elements 96
- condensation 34–35, 36, 37, 52
- condensation flow 137
- condensation sequence 35, 215–216, 217
- conduction 62
- conservation of angular momentum 28, 29–30, 62, 250
- continental crust 50
- continental flood basalts 100–101
- continuous ejecta 73
- convection (in an atmosphere) 134–135, 189, 194
- convection (within a planetary body) 50, 55, 62, 64, 175
- convective ascent region (within an eruption column) 103
- Copernicus (lunar crater) 73
- Copernicus, Nicholas 10
- core (of a planetary body) 49, 50, 53, 55, 58, 173, 176, 251, 252
- core formation 59–60, 173
- Coriolis effect 135–136, 138, 189–190, 192, 194
- corundum 33
- cosmic ray exposure age 229–230
- cosmic rays 218, 229
- cosmic sediments 212
- cosmic spherules 207
- crater production flux 84
- cratering time-scale 83–91, 162
- craters (impact) (see also dark halo craters) 42, 69–91
- crust (of a planetary body) 49, 50, 53, 55, 252
- cryovolcanism 96, 98

- Dark Spot (Neptune) 193–194, 202
- dark halo craters 110
- Deimos 16, 45
- dense (interstellar) cloud 217
- density (of planetary bodies) 6–7, 15–16, 21, 41, 44, 45
- density (of solar nebula) 36, 37
- deuterium (^2H) 161
- diamond 218, 257
- diatomic molecules 122, 123, 254
- differentiation (of a planetary body) 52–53, 54, 173, 175, 220–221, 256
- discontinuous ejecta 73
- dissociation 141
- dissolved gas (in a magma) 55, 102–103
- diurnal variation (in an atmosphere) 124, 131, 134, 136, 137
- dust (grains) 27, 34–38, 44, 157, 212, 215, 233–236, 258

- Earth 31, 39, 41, 48–51, 61, 108–109, 129, 131, 143–149, 162–165, 226
- Earth-crossing asteroids 18, 207
- ecliptic plane 12, 28
- EET A79001 (meteorite) 225–232
- effective temperature 127–129, 173
- effusive eruptions 99–102
- ejecta (from impact craters) 72, 73–74
- ejecta (volcanic) 104–105
- electric dipole 122–123
- electromagnetic induction 222
- electron pair bond 141–142
- Enceladus 98
- endothermic reaction 140, 254
- energy barrier 140
- equilibrium (see also chemical equilibrium) 132–133, 156
- equilibrium constant 184
- Eros 13
- eruption columns 103–104, 105, 106–107
- eruption rate 100
- escape speed (escape velocity) 22–23, 157, 197, 227
- ethane (C_2H_6) 179–181, 198–199, 256
- ethyne (C_2H_2) 178–181, 198–199, 256
- eukaryotic cells 163–164
- Europa 16, 44, 98
- evaporation sequence 35
- exothermic reaction 141, 254
- explosive eruptions 99, 102–107

- falls (meteorites) 204, 257
- feldspar 35, 221, 227
- finds (meteorites) 204
- fire fountains 102

- ul style="list-style-type: none; padding-left: 0;">
- fireball 208
- flood basalt 100–101, 113
- focus (of an ellipse) 12
- fragmentation 36, 42
- fusion crust 207
-
- galilean satellites 44–45
- Galileo Galilei 10, 44
- Galileo spacecraft 16
- Ganymede 16, 44
- gas (in magmas/volcanoes) 102–3, 107, 117
- gas (in the solar nebula) 36, 40, 41, 44
- gas chromatography 123, 254
- gas thrust region 103
- Gaspra 16
- gegenschein 233
- giant impact 54–55, 57–58, 251
- giant planets (see also Jupiter, Neptune, Saturn, Uranus) 16, 29, 40–41, 42, 52, 169–170, 182–183, 250
- Giotto spacecraft 233
- glass 73
- grains – see dust
- gravitational attraction 35, 38, 39, 40
- gravitational collapse/contraction 27, 28
- gravitational energy 28, 34, 45, 54
- gravitational focusing 38, 40
- Great Red Spot (Jupiter) 190–191, 193, 202
- Great White Spot (Saturn) 191
- greenhouse effect (see also runaway greenhouse effect) 129–130, 160, 255
-
- ^2H – see deuterium
- Hadley cell 134–136
- half-life 60, 222
- Halley's comet – see comet Halley
- haze 187, 197, 200
- heat excess 175, 256
- heat sources (in planetary bodies) 53–55, 58, 59–62, 175–176, 222
- heat sources (in the solar nebula) 34
- helium 40, 41, 170–171, 178–180, 250, 255
- heterogeneous accretion 51–52
- heteronuclear molecules 123
- homogeneous accretion 52
- homolytic dissociation 141
- homonuclear molecules 123
- hot-spot volcanism – see advection
- hydrated minerals 35, 216, 235
- hydrazine (N_2H_4) 197
- hydrocarbons (see also under individual names) 181, 187, 198–200, 216
- hydrodynamic escape 231
- hydrogen 40, 41, 155–157, 173, 198–200, 250, 255
- hydrogen sulphide (H_2S) 185
- hydroperoxyl radical (HO_2) 152–153, 255
- hydrothermal alteration 211, 216
- hydroxyl radical (OH) 141, 145–146, 152–153, 254
- hyper-velocity impacts 72
- Hyperion 45
-
- ice 34, 36, 37, 41, 44, 45, 52, 53, 64, 77–78, 96, 97–98, 169, 181, 250, 255–256
- impact cratering (see also craters, hyper-velocity impacts) 69, 226–288
- impact melt 73
- impact flux 84, 90
- incompatible elements 96
- infrared spectroscopy – see vibrational spectroscopy
- infrared transmission spectra 234–235
- inner core (of Earth) 50
- interiors of planets 48–53, 62–64, 170–173
- interplanetary dust (see also dust) 233
- interstellar medium 217
- Io 16, 44, 61, 115–117, 226, 254
- ionic substances 34
- ionization 131
- iron 19, 33, 48–51, 55, 57, 96
- iron meteorites 205, 220–221, 223–224
- iron sulphide 33, 52, 215
- isopachs 105
- isopleths 105
- isotope compositions – see stable isotope compositions
- isotope exchange 215, 231

- isotope fractionation 213, 229
- jet stream 135
- Jupiter 16–17, 36, 40, 42, 45, 170, 173–175, 185, 189–191
- Juvinas (meteorite) 220–223

- Karla (1993 FW) 42, 45
- Kepler's second law 13
- Kepler's third law 13
- Kepler's first law 11
- Kepler, Johannes 10
- kernel 40, 42, 45, 52, 169, 173, 176, 250, 256
- kinetic energy (see also thermal kinetic energy) 42, 54, 59, 102
- Kirkwood gaps 17
- Kodaikanal (meteorite) 257
- komatiite 96, 254

- latent heat of fusion 98
- lava (see also basalt, flood basalt) 48, 65, 80
- life (geological record of) 62, 163–164
- lithosphere 63, 64, 77, 78, 80, 95, 99, 252
- long-period comets 20
- low-speed layer 64

- ^{26}Mg – see magnesium-26
- M-type asteroids 209
- Magellan spacecraft 110
- magma 65, 95, 220–222
- magma ocean 55, 59, 159
- magnesium 33, 96
- magnesium-26 223
- magnetic dipole 50, 171–172, 176, 251
- magnetic field 50, 188–189, 255
- main sequence stars 27
- mantle (of a planetary body) 49, 50, 53, 55, 58, 252
- mare (maria) 80, 84
- Mars 39, 45, 89–91, 112–115, 134, 136, 151–153, 161–162, 225–232, 253, 254
- Mars Observer mission 112
- Martian atmosphere 230–231
- maskelynite 227
- mass extinctions 69, 71, 100
- mass spectrometer 213, 223, 233
- mass spectrometry 123–124, 254
- mass-loss (from solar nebula) 31
- massive solar nebula model 29, 31
- matrix (of a meteorite) 37, 212, 215, 236, 257
- Mercury (planet) 39, 110, 121, 158, 226, 253
- mercury (element) 174, 255
- mesosphere 131
- metallic bonding 173–174
- metallic hydrogen 173–175
- Meteor Crater 71
- meteor stream 209
- meteorite showers 204, 210, 223
- meteorites (see also under individual names) 19, 31, 34, 37, 38, 50, 54, 204–207, 208–209
- meteoroids 207–208, 226
- meteors 19, 208–209
- methane (CH_4) 33, 45, 97, 145–146, 165, 173, 178, 179, 198–199, 201, 256
- micrometeorites 206, 234
- micrometeoroids 207, 209, 233
- mid-ocean ridge basalts (MORB) 96, 253
- Mimas 16
- minerals 33, 35, 50
- minimum solar nebula model 29, 31
- Moho 50
- molecular substances 34
- Moon 16, 56–58, 70–71, 73–75, 78–81, 83–88, 109–110, 209, 226, 233, 251
- MORB – see mid-ocean ridge basalts
- multi-ringed basins 80
- Murchison (meteorite) 210–211, 215–219

- nakhilites (see also SNC meteorites) 225
- nebular theory 27–29
- neon 256
- Neptune (see also Dark Spot) 40, 45, 175–176, 191–194, 255
- neutrons 212
- nickel 48–51, 57
- nickel-iron 52, 206, 224
- nitrogen 45, 123, 125, 181–182, 228–230, 233

- nitrogen cycle 146–147
- nitrogen fixation 146–147
- nitrogen isotopes 229–230
- noble gases (see also radiogenic noble gases) 122, 156, 158–159
- normal distribution 83–84
- nova 214
- nucleus (of a comet) 20, 21
- number density 141

- oceanic crust 50–51
- oceans 126, 163
- olivine 35, 50, 215, 227, 234–235
- Olympus Mons (Mars) 113
- Oort cloud 22, 42
- orbital eccentricity 12
- orbital inclination 12
- orbital period 13
- orbital resonance 17–18, 61, 251
- ordinary chondrites 206, 209
- organic compounds 216–217, 233, 236
- Oriente basin (on Moon) 80, 252
- outer core (of Earth) 50, 52
- outgassing 55, 144, 150, 156, 158, 159, 255
- oxidation 145
- oxides (convention in rock composition) 50
- oxidized substance 145
- oxidizing substance 145, 153, 254, 255
- oxygen 33, 50–51, 145
- oxygen isotopes 56, 212–213
- oxygen isotope compositions 212–215, 226, 231, 257, 258
- ozone 131, 148, 164, 255

- P-waves 49–50
- palimpsests 78
- parallax – see trigonometric parallax
- partial melting 53, 63, 96, 97, 220
- partial pressure 131–133
- peak-ring basins 80
- peridotite 50
- perihelion 8, 209, 233

- permafrost (on Mars) 151
- Phobos 16, 45
- phosphine (PH₃) 178, 179
- photochemical smog 200
- photochemistry 142, 198–200
- photodissociation 142, 149–151, 160, 163, 197, 198–199, 231, 255
- photosynthesis 143–144, 163
- planetary embryos 38–40, 42, 45, 52, 53–54, 57–58
- planetary satellites 43–45
- planetesimals 38–39, 40, 42, 43, 44, 52–54, 250
- plate recycling 65, 113, 160
- plate tectonics 63
- Pluto 12, 45
- polar ice-caps 125–126, 151–152
- potassium 48, 50, 95, 96, 158, 255
- potential energy 140
- pre-solar materials 212, 214–215, 218
- prebiotic atmosphere (of Earth) 163–164
- precipitation 126, 133
- pressure (in solar nebula) 35
- primitive magmas 96
- primitive meteorites 31, 210–211, 236
- prograde orbits 28, 44
- prograde spin 28
- protons 212
- protosatellite disc 44–45, 52
- protoSun 28–31, 34
- pyroclastic eruptions 99, 102–107, 109–110, 113–114, 116
- pyroclastic rocks 99
- pyroclastic flows 105–106, 109
- pyroxene 33, 35, 50, 52, 221, 227, 234–236

- radiation (as a means of heat transfer) 28, 34, 62
- radiation pressure 234
- radicals 141, 197
- radiogenic heating 54, 60–61, 98, 175, 222–223, 252
- radiogenic noble gases 158, 255
- radiometric dating 60–61, 84, 223, 226
- rain – see precipitation
- raisin-bread structure 224

- rampart craters 76–77
- rare-earth elements (REE) 221–222
- rarefaction wave 72
- rate of reaction 141
- rays (of ejecta) 73
- red giant 218
- reduced substance 145, 180
- reducing atmospheres 180
- reducing substance 145
- reduction 145
- refractory substances 35, 57, 251
- relative molecular mass 123
- retrograde orbits 45
- rings (around planets) 17, 45
- rock 33, 37, 41, 44, 45, 250
- rotation period (of giant planets) 188–189
- rotational spectroscopy 122
- runaway greenhouse effect 160, 165, 255
- runaway growth 38–42, 54

- s- (or slow) process (reaction) 218
- S-type asteroids 19, 209
- S-waves 49–50
- satellites (of planets) 36, 43–45, 52, 251
- saturation (cratering) 84
- saturation vapour pressure diagrams 132–133
- Saturn 40–41, 45, 170–171, 173–175, 191
- seasons 136
- secondary craters 73, 84, 252
- seismic waves 49–50
- selection rules 123
- self-compression 41, 48, 171, 176, 253, 255
- semi-major axis (of an ellipse) 12
- shergottites (see also SNC meteorites) 225
- shock waves 72–73
- short-period comets 20
- SiC grains 218–219
- siderophile elements 57–58, 251
- silicates 33, 34, 37, 52–53, 96, 173
- silicon 33
- silicon carbide 218–219, 236
- sinks (for atmospheric components) 143, 144, 150, 152
- size-frequency distribution 83–90, 252
- Smiley (1993 QB₁) 42, 45
- SNC meteorites 226–232, 257–258
- Solar System (nature of) 10–23
- Solar System (origin of) 27–31
- solar nebula 27–37, 40, 41, 44, 51, 210, 212, 214–218, 236
- solar radiation 131, 149, 151, 231
- solar wind 255
- solid-state convection 62–64
- sources (for atmospheric components) 143, 144, 150, 152
- spallation 218
- specific heat capacity 97–98
- spectroscopy 31
- stable isotope compositions 212–214, 217, 229, 231, 233, 257
- stable isotopes 56, 212
- steady state (atmospheric composition) 139, 149
- stellar winds 31, 38
- stony meteorites 205
- stony-iron meteorites 205
- stratosphere 131, 147–149
- stromatolites 163
- sulphur 50, 117
- sulphur cycle (Venus) 149–150
- sulphur dioxide 117
- sulphuric acid (H₂SO₄) 126, 134, 149, 255
- Sun (see also protoSun and T Tauri phase) 29, 31–32, 38, 210, 222
- super-rotation 136
- supernova 54, 214
- surface heat flow 115
- surface reflection 126
- surface temperatures 128–129, 136, 149, 254
- synchronous rotation 61

- T Tauri phase (of the Sun) 31, 39, 41, 158, 159, 222
- T Tauri stars 27, 34, 38
- T Tauri wind 38, 41, 55, 250
- tail (of a comet) 21
- tektites, 73
- temperature (in a planetary body) (see also effective temperature) 53–54, 97

- temperature (in a protosatellite disc) 45
- temperature (in the solar nebula) 34–35, 36, 37, 52
- tequila sunrise 175, 242
- terminal fall speed 104–105, 253
- terraces (within craters) 80
- terrestrial fractionation line (TFL) 214, 226, 257
- terrestrial planets 16, 39, 52, 250
- TFL – see terrestrial fractionation line
- Tharsis volcanic province (Mars) 113
- thermal escape of gases 23, 157, 160, 162, 255
- thermal kinetic energy 28, 34, 45
- thermal tide 137
- thermosphere 130–131, 183
- third body 142, 197, 254
- three-body collisions 142, 197
- tidal dissipation – see tidal heating
- tidal forces 45, 98
- tidal heating 61–62, 97–98, 252, 253
- Titan 181–182, 196–200, 201
- Toutatis 42
- trade winds 135
- trigonometric parallax 13
- Triton 43, 45, 254
- troilite 33, 52, 215
- troposphere 130, 143–147, 183
- Tunguska 21
- turbulence (in the solar nebula) 30, 35, 38
- Tycho (lunar crater) 73

- ultraviolet flux (of the Sun) 231
- umbrella region 103
- Uranus 40, 45, 175–176, 194–195, 255

- Valhalla basin (on Callisto) 252
- Vega 27
- Venus 39, 61, 110–112, 129, 133–134, 136, 149–151, 159–161, 226, 251, 253, 254
- Vesta 24, 223, 249, 257
- vibrational spectroscopy 122–123, 178, 254, 256
- vibrationally excited molecules 131
- Viking (Mars missions) 124, 228–230
- viscosity (of magma) 99–100, 101–102
- viscous drag (in the solar nebula) 30–31, 34, 38, 250
- viscous relaxation (of topography) 78
- volatile inventories 127, 155, 159, 161
- volatiles (volatile substances) 35, 57–58, 102–103, 106–107, 117, 155, 210, 211, 253
- volcanic ash 99
- volcanism 65, 159
- volcanoes 48, 159

- water (see also ice) 33, 34, 36, 37, 155–156, 159–163, 173, 212, 256
- weathering 145
- Widmanstätten pattern 224
- wind 135–137
- wind speed 135–136, 187–189, 254, 256

- yield strength 102

- zodiacal light 233
- zones 189, 191